UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

THE DEVELOPMENT OF A SCALE TO MEASURE PERCEPTIONS OF THE ADVANCED AUTOMATED AIRCRAFT TRAINING CLIMATE

by

PREVENDREN NAIDOO

Submitted in partial fulfilment of the requirements for the degree

PHILOSOPHIAE DOCTOR

(ORGANISATIONAL BEHAVIOUR)

in the

FACULTY OF ECONOMICS AND MANAGEMENT SCIENCES

at the

UNIVERSITY OF PRETORIA

APRIL 2012

Supervisor: Professor Dr L. P. Vermeulen

Co-supervisor: Professor Dr P. Schaap

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

FACULTY OF ECONOMIC AND MANAGEMENT SCIENCES

Declaration

I, Prevendren Naidoo,

Declare the following:

1. I understand what plagiarism entails and I am aware of the University's policy in this regard.
2. I declare that the thesis entitled *The development of a scale to measure perceptions of the advanced automated aircraft training climate* is my own original work, both in content and execution. Where someone else's work was used (whether from a printed source, the World Wide Web or any other paper or electronic source), due acknowledgement was given and reference was made according to departmental requirements. The Harvard method was used to acknowledge all references.
3. I did not copy and paste any information directly from an electronic source (e.g. a web page, electronic journal article or compact or digital video disc) into this document.
4. Apart from the guidance and support from my study leaders, I did not make use of another person's previous work or submit it as my own.
5. I did not allow and will not allow anyone to copy my work with the intention of presenting it as his/her own work.

April 2012

Signature

Date

ACKNOWLEDGEMENTS

Completing a study of this magnitude cannot be achieved without the selfless support, encouragement and help of others. I wish to express my sincere appreciation and gratitude to the following:

- First and foremost, my supervisor, Professor Dr Leo Vermeulen, for taking me under his “wing” more than 10 years ago. His guidance and passion for aviation was priceless in inspiring this research project. Without his continued support, patience, and depth of knowledge in the topic, this study would have never succeeded.

- My co-supervisor, Professor Dr Pieter Schaap, for his wisdom, advice and encouragement.

- My family and friends for their continuous support and encouragement during some of my most difficult and trying times as an academic hermit.

- Mrs Idette Noomé, for her excellent advice and patience in editing the thesis.

- Ms Rina Owen at the University of Pretoria’s Department of Statistics, for her support with, and knowledge about, the more complex computational problems I encountered in this study.

- The pilots, academics and experts in the industry who unselfishly gave up their time to provide the constructive feedback and critique needed to make this study a success. The objectives of the research would never have been reached without the support of this truly wonderful group of people.

- Finally, if it were not for the individuals who sacrificed both their lives and many thousands of hours perfecting a magnificent machine in order to fulfil the elusive dream of heavier-than-air flight, this thesis would never exist.

SUMMARY

The development of a scale to measure perceptions
of the advanced automated aircraft training climate

by

PREVENDREN NAIDOO

SUPERVISOR: Professor Dr L. P. Vermeulen

CO-SUPERVISOR: Professor Dr P. Schaap

DEPARTMENT: Department of Human Resource Management

DEGREE: Philosophiae Doctor (Organisational Behaviour)

Commercial air travel is regarded as the safest mode of transportation known to humankind; however, every year people lose their lives from aircraft accidents and incidents. In addition, the financial impact of an air disaster can destroy an airline organisation. Studies have found that in adverse events involving highly advanced aircraft employing complex automation, human factor issues, and particularly pilot training, continue to play a significant causal role. Special attention should therefore be paid to the *training* of airline pilots, who are ultimately the last line of defence in aircraft operations. Airline pilots' perceptions of the training climate associated with advanced aircraft can be a pervasive and powerful determinant of training outcomes and eventual flight deck behaviour.

The study undertook to develop a valid and reliable instrument to measure airline pilots' perceptions of the training climate associated with advanced aircraft equipped with highly complex automation. The goal was to construct a questionnaire by operationalizing an unobserved hypothesised construct (perceptions of the advanced automated aircraft training climate) based on three levels of analysis (the microsphere, mesosphere and macrosphere). The study also attempted to explore the statistical relationship between the demographic variables of the respondents and the latent factors of the construct.

In order to meet the research objectives, the study began with a thorough review of the current literature on the topic to develop a systems model of the main construct under investigation. The review included a critique of the theory on organisational climate, learning, training and education, of historical data on aircraft automation, of human factors, and of aircraft accident investigation principles and case studies. The objectives of the research were fulfilled by strictly observing a positivist paradigm, and engaging in a quantitative exploration, triangulating methods with data captured from a purposive sample of the target population. The empirical study was completed in four phases. Firstly, the research construct was operationalized and the items in the proposed questionnaire validated by a panel of subject matter experts using Lawshe's (1975) content validity ratio (CVR) technique. Inter-rater bias was assessed using Cochran's Q test. This application resulted in the retention of 42 items. Secondly, factor analysis and item analysis was performed on the responses of the respondents for the development of the final 33 item measurement instrument. Thirdly, to explore the relationship between the demographic variables and the latent factors of the main construct, an appropriate non-parametric family of statistics was selected to gain a deeper understanding of the phenomena associated with the data. Finally, a logistic regression analysis that included specific demographic variables was performed for the development of a model to predict a pilot's perception of the training climate associated with advanced automated aircraft.

A non-probability purposive sample of 17 subject matter experts and 229 qualified South African airline pilots was used to accomplish the goals of the study. The underlying structure of the advanced Automated Aircraft Training Climate Questionnaire (AATC-Q) was derived from the results of a Principal Axis Factor (PAF) analysis using a promax (Kappa-4) rotation. The number of factors extracted from the data set was based on a modified version of Horn's (1965) parallel analysis, namely the Monte Carlo simulation algorithm designed by O'Connor (2000). Three core factors explained most of the underlying variability in the main construct. The first factor was a composite at the macro and meso levels of analysis, whilst the second and third factors became fragmented at the micro level of analysis. These three factors were then labelled *Organisational Professionalism*, *Intrinsic Motivation* and *Individual Control of Training Outcomes*. The quality and rigour of the derived scale were demonstrated by its content and construct validity. Overall, satisfactory results from computing Cronbach's coefficient alpha showed that the measurement scale was also reliable.

The effect of the demographic variables on airline pilots' perceptions of the advanced automated aircraft training climate was determined by computing relationships and comparing the responses from different categorised subsets with one another, by means of a non-parametric MANOVA and non-parametric analysis of variance. The results of these tests revealed that *Flight Deck Position, Size of the Airline, Computer Literacy* and *Flight Experience* had a significant effect on a pilot's perception of the training climate. Results from a logistic regression model indicated that the interaction between pilots' experiences and their perceived level of computer literacy (on a sigmoid curve), their actual experience in advanced aircraft, and their preferences for route and simulator training, were related to whether a pilot perceived the advanced aircraft training climate as favourable or not. The overall percentage of cases for which the dependent variable was correctly predicted by the regression model was computed at 63.8%.

This study represents a vital step toward an understanding of the dimensionality of the learning, education and training for, and the actual operation of, highly advanced commercial aircraft, which employ complex automation. The results provide sufficient empirical evidence to suggest that the research findings may be of particular interest to aviation psychologists, aviation safety practitioners, and airlines engaged in training pilots to operate advanced aircraft.

Keywords: automation in aviation, advanced aircraft, advanced aircraft training, Automated Aircraft Training Climate Questionnaire (AATC-Q), aviation training, aviation exploratory study, human factors, Individual Control of Training Outcomes, Intrinsic Motivation, measurement scale, Organisational Professionalism, perceptions of aviation training, training climate.

CONTENTS

CHAPTER THREE: LITERATURE STUDY – THE ADVANCED AIRCRAFT TRAINING CLIMATE

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE:

INTRODUCTION –
THE PROBLEM AND ITS CONTEXT

1.1 INTRODUCTION

On 14 October 1947, a young pilot, Captain Charles E. ("Chuck") Yeager, broke the sound barrier. Later, he said: "I was always afraid of dying, always. It was my fear that made me learn everything I could about my airplane" (Yeager, cited in Harrison, 2000:96, own emphasis). This statement encapsulates the profound importance of training, even for some of the most famous names in aviation circles. In order to master the machines they fly and so change the course of history, it was critical that these pilots gained a deep understanding and knowledge of their aircraft. Modern airline pilots operating some of the most advanced machinery known to humankind are no exception to this rule.

Today's successful airline organisations, which operate advanced commercial aircraft employing highly complex automation, have consistently found that well-trained pilots are the cornerstone of their profits and ultimate business survival. The situation is complicated by the fact that, in order to remain competitive in an industry renowned for failure and bankruptcy, airline companies are also obliged to invest in expensive modern and more efficient aircraft. This investment includes training highly competent pilots (Australian Transport Safety Bureau, 2007; Walters, 2002). Although not all training situations are as dramatic as Yeager's remark implies, training human beings to handle an advanced automated aircraft can be an extremely expensive, challenging and time-consuming exercise (Johnston, Fuller & McDonald, 1995) that require effective organisational practices and structures, dedicated and skilful instructors, and most importantly, motivated learners (Telfer & Moore, 1997).

Air travel in commercial advanced aircraft is currently rated as the safest mode of transportation (Boeing, 2009); however, people still continue to perish as a result of aircraft accidents, and the financial impact of such accidents can close down an airline company (FAA, 1996). The consequences of an aircraft accident are devastating, for not only the company involved, but also the communities that such an event ultimately affects. Alarmingly, the reported probable cause cited in over half of aircraft crash investigations is pilot or human error (Billings, 1997; Cockburn, 2007; NTSB, 2009). Studies have shown that in highly advanced automated aircraft accidents, human factor issues in general, and pilot training in particular, often play a significant role (Ishidi & Kanda, 1999; Kaminski-Morrow, 2009; Rouse & Morris, 1987; Sarter, 1996).

Moreover, O'Hare, Wiggins, Batt and Morrison (1994) have found that, apart from a multitude of variables responsible for a significant portion of the failures in the system, specific attention must be paid to the training received by the pilots who were at the controls of the aircraft. In a quantitative study, Sarter (1996) concluded that a thorough understanding of the training environment can assist researchers in identifying and mitigating at least 40% of the human factor variables associated with aircraft accidents. In a similar study, Machin and Fogarty (2003) found specific interactions between individual variables, training methods and situational factors that influence learning outcomes. These conclusions point to psychological reasons for the root causes of success or failure in a transfer of knowledge. In fact, Telfer and Moore (1997) found that incongruity between the pilots, the machines and the organisation, could explain some of the latent systemic training problems in an airline. Desler (2002) suggests that a systemic nature of a psychological climate, implies aligning three domains of organisational behaviour, that is, at an individual, group and organisational level of analysis.

The aforementioned introduction positions the existence of an ongoing need for further scientific scrutiny to determine the factors that may have an impact on the success or failure of pilot training, related to the safe operation of advanced automated aircraft. Only from scientific analyses and a genuine need to ultimately understand the human-technology dyad, can these human factor issues be effectively uncovered and thereby addressed in the aviation industry (Johnston et al., 1995).

1.2 THE RESEARCH CONTEXT

According to Boeing (2010), the overall volume of commercial air traffic will increase threefold over the next 20 years. However, it is expected that the commercial aircraft accident rate will decrease steadily, in major part due to the exponential advances in technology (Boeing, 2009; CAA, 2011; FAA, 1996). Both regulators and manufacturers are of the opinion that increasing technology in aircraft will reduce human input, with a subsequent reduction in human error, thereby mitigating the total accident rate over time (Bainbridge, 1983; Wald, 2009). The proliferation of highly automated flight decks and the increased use of computer-based heuristics in aircraft have reduced pilots' workload and have, to a degree, eliminated adverse aspects of the human element (Funk & Lyall, 2000). However, some experts in the industry tend to disagree with such assessments of the situation (Barker, 2011; Bent, 1996).

Degani, Shafto and Kirlik (1995) have argued some time ago, that the rapid increase in computer technology predicted for the future, with its subsequent impact on advanced automated aircraft, will have a profoundly negative effect on the human-machine dynamic. Their argument is based on the premise that, the resultant effect of an increase in technological complexity implies that overall, far superior human cognitive effort is required to manage the new technology, thereby increasing the likelihood of human error. More than twenty years ago, Bainbridge (1983) began pointing out a paradox. It was hypothesised that the more complex a system becomes, the more critical or important the contribution by the human operator (Bainbridge, 1983). This paradox has begun manifesting within the last decade, highlighting the need for a deep understanding of technology and its impact on human behaviour (Cockburn, 2007; Hradecky, 2011).

The intense cognitive requirements needed to operate advanced aircraft have resulted in a multitude of automation and computerisation debates (Barker, 2011; Poprawa, 2011; Bent, 1996; Parasuraman & Riley, 1997). Although increased computerisation of aircraft may promise an improvement in their operation (Ausink &

Marken, 2005), Funk and Lyall (2000) have argued that too much automation can also prove disastrous. Hence, the inherent dichotomy between humans and advanced machines implies that an improvement in learning methodologies for new technology may require a paradigm shift in future organisational training activities (Telfer & Moore, 1997). It is therefore increasingly important to research training for the users of technologically advanced systems.

Research conducted in the field of aircraft automation over the last decade has revealed how critical latent training and related safety issues in the human-machine system are (Walters, 2002; Wickens, 2000; Wiener, 1998). A fatal combination of limited technical knowledge and its incorrect application in both normal and emergency situations was found to be a significant contributor in technology-related aircraft accidents (Baum, Gatchel & Schaeffer, 1983), and more recently reported to be a significant contributor to accidents involving the latest technologically advanced commercial aircraft (Cockburn, 2007; Hradecky, 2011).

An illustration of the complexities of aircraft technical know-how and its application is provided by Parasuraman and Riley (1997), who have determined that the intricacy of complex systems (for example, automatic thrust levels, computer mode changes and so forth) in both normal and abnormal flight situations can distract a crew to such an extent that the basic management of the overall aircraft may be jeopardised. These consequences are directly linked to inappropriate usage of the autopilot system (Poprawa, 2011). Operating an advanced aircraft system correctly is necessary so as to harness safety; which was the original intention of the advantages found in technology. For instance, zero visibility landings can only be safely accomplished by automation, and therefore the correct and appropriate use of the autopilot is mandatory in adverse weather conditions and highly recommended whenever visual conditions fall below 1KM (South African Airways, 2007). This is an example of the safety benefits associated with the correct use of automation technology. Researching and understanding why and how the incorrect use of technology can profoundly compromise flight safety is an important requirement in accident prevention and mitigation efforts (Ausink & Marken, 2005).

According to Bent (1996) and Sherman (1997), it is not necessary for a pilot to know all the intricate details of precisely how an aircraft's technology is designed and built in order to operate the aircraft safely. This is one reason why specific licences are stipulated by the aviation regulator for particular facets of aircraft operation (CAA, 2011). For example, an engineer is only qualified to build an aircraft to a certain degree of specification, but may not have the skill to fly it safely with the knowledge acquired in attaining that qualification. Similarly, a pilot is not required to have the technical depth of knowledge of an engineer to operate the aircraft safely. Research can help bridge the gap between the two areas (that is, flying skill versus in depth technical knowledge) to enhance understanding of the nature of accidents related to a breakdown in knowledge between design and operation. For instance, Sherman's (1997) research explored automation training, focusing on airline pilots in the United States, and established significant differences between the ways different categories of respondents view various advanced automated aircraft types. However, his study did not address all the relevant issues – airline pilots' perceptions of the training climate associated with an advanced automated aircraft were not analysed.

Telfer and Moore (1997) diagnosed the training climate associated with the general aviation sector. They deduced that there is an inherent need for aviation organisations to develop more effective learning philosophies and methodologies proactively in order to reduce human training-related incidents, and thereby improve overall flight safety. Flight safety also has a basic financial imperative. The link between flight safety and an improved bottom line is emphasised in a comment such as "safety makes good business sense" (Walters, 2002:110). Investment in flight training efforts, such as research regarding learning to fly new technology aircraft, can thus add value for the organisation as a whole. The benefits of the leveraging effect of training to an airline are immense. Indeed, Abraham (1990:20) posits that the management of aviation safety by systemic improvements is so basic and fundamental that "profit, pride and politics [should] take a back seat to…efforts devoted to safety".

A preliminary review of the scholarly database at the start of the study reported here, suggested that very little empirical research had previously been conducted on the topic covered in the current research topic, namely, the development of a scale to measure perceptions of the climate associated with advanced automated aircraft training. Hence, a gap in the knowledge of aviation human factors was uncovered. Where any prior analyses of how current airline pilots perceive learning for flying technologically advanced aircraft have been conducted, these analyses were inconclusive regarding the psychological attributes of trainee pilots (Ausink & Marken, 2005). An exploration of this subject and the challenges encountered after analysing this specific area, has ultimately led to a deeper understanding of current commercial aviation human factor issues.

In determining empirically the value associated with safety, Gegax, Gerking and Schulze (1991) assert that no monetary saving should be placed above the advancement of safety elements affecting both the organisation and, in particular, the flying public. Their study highlights the importance of more research into both overt and latent aviation safety constructs, and particularly aviation training. Governments, researchers, and economists have found it very difficult to quantify the cost of a commercial aircraft accident, specifically where hundreds of lives are lost (International Civil Aviation Organisation, 2001).

1.3 THE RESEARCH PROBLEM AND ITS SIGNIFICANCE

Creswell (2002) argues that the significance of a study can be assessed by both its scholarly contribution and the improvements it brings about in policy and practice.

There may be some truth in Orben’s quip that “[t]o err is human and to blame it on a computer is even more so” (cited in Zametti, 2008:128). Shifting the blame has long been a common feature of reports on aircraft accidents, and computers have become a new target for blame (Kaminski-Morrow, 2009). As long as manufacturers continue to improve aircraft by designing and installing ever more advanced computer-based systems, there will always be the related human factor issues (Baum *et al*., 1983), even when it seems as if manufacturers are attempting to design out the human factor (Barker, 2011), by limiting the level of control apportioned to the human operator (Wickens, 2000).

Despite ongoing research over the last two decades concerning flight deck automation relating to and affecting safety management systems, problems with training activities related to the implementation of digital automation continue to be reported (Cockburn, 2007; Kaminski-Morrow, 2009; Mitchell, Vermeulen & Naidoo, 2009; Wiener, 1993). Preliminary investigations into the accident in early 2009 involving a Turkish Airlines Boeing 737-800 suggests that (inadequate) pilot training may have contributed significantly to the fatal crash (Dutch Safety Board, 2009). Subsequently, Airbus Industries and the Boeing Company have issued an advisory to all crew operating their aircraft to be vigilant in monitoring automation systems such as auto-thrust, and to maintain control during wing stall situations (a very basic aerodynamic concept taught to pilots at flight school). To rectify various short-term issues, aircraft manufacturers will send out a notice to their customers to implement certain changes to flight operations or modify equipment. A bulletin from Airbus clearly states, “during all phases of flight, flight crew must monitor and crosscheck all primary flight parameters and the FMA” (Airbus, 2011b). The Flight Mode Annunciator or FMA is a primary instrument used by the pilots to monitor the status of the flight control computers. In addition, this specific notice was colour coded red, which indicates a critical requirement. This has served to highlight the fact that manufacturers are becoming more concerned with the recent trend in critical weak areas relating to human factors in the design, monitoring and operation of advanced computerised aircraft. In fact, it was observed that the two operations bulletins found in the Airbus A330-200 Quick Reference Handbook (QRH) on board the flight deck, are both related to flight automation systems.

Scientific interest in a training climate construct associated with the operation of advanced aircraft has arisen as a result of the paradigm shift brought about by the rapid technological advances in aviation (Telfer & Moore, 1997). Transition training (transferring from an analogue aircraft to a digital aircraft) is often cited as a primary concern when analysing human factor issues related to the actual *comprehension* of an advanced digital aircraft system (James, *et al.,* 1991; Naidoo, 2008; Wise, *et al.,* 1994). The systemic nature of flight operations, compounded by a complex environment, leaves room for human operator errors or lapses which may have a catastrophic outcome (Ishida & Kanda, 1999; Parasuraman & Riley, 1997; Sarter, 1996). A review of various analyses of the latent structure of pilots’ perceptions of advanced automated flight decks shows that errors in human behaviour related to training issues can have a significant impact (Naidoo, 2008; Wiener, 1993). Hence, training has played, and will continue to play, a vital role in how airline pilots perceive advanced flight deck automation, which in turn directly influences levels of safety. According to the National Transportation and Safety Board (NTSB, 2009), a significant proportion of accidents involving highly modernised commercial jet airliners can be attributed to poor training. Therefore, efforts to reduce this contributory factor will have a direct and observable impact on safety. Furthermore, Risukhin (2001) suggests that training through higher order learning systems is regarded as an antecedent to a positive perception of advanced technology aircraft computerisation, which can, in turn, enhance the training environment.

Quantitative research into an organisation's *training climate* helps scientists understand the multitude of variables responsible for employees' pedagogical development and the final achievement of the overall learning objectives (Naidoo, 2008; Tracey & Tews, 2005). Collecting data from psychologically measurable variables in terms of the advanced aircraft training climate will effectively create the conduit necessary for developing an accurate construct for the aviation industry. A problem in aviation is the failure to develop accurate and measurable constructs (Sherman, 1997). For instance, Desler (2002) points out that when considering the accuracy of the design of an aviation related training climate construct, it is important to understand details such as the resources, practices and priorities of the organisation in terms of the instructor, facilitator, and more importantly, the attitudes, motivation, strategies and learning styles of the student.

To address the extent of the issue of insufficient measurable constructs within the aviation industry, previous theory, case studies and reports were consulted. Constructs and concepts from organisational behaviour have been operationalized in this research in order to link up with training-related attitudes and outcomes. The premise is that only effective training methodology can significantly influence flight safety from a learning point of view. Singh, Sharma and Singh (2005) found the empirical evidence to support this statement by showing that longer training periods ($M=41.67$, $SD=5.51$) were only marginally better in changing performance than shorter periods were ($M=37.82$, $SD=5.51$). Such experiments demonstrate the current research problem practically, by accurately and empirically substantiating the role of quality in the training and learning environment. Quality, and specifically the perception of training quality, can ultimately impact behaviour of pilots in the aircraft itself (Naidoo, 2008).

The theory of planned behaviour surmises that the quality of *learning* can directly influence final or actual outcome behaviour (Fishbein & Ajzen, 2001, Meister, 1999). Therefore, the development of a measurement of this environment potentially makes the current study a significant contributor to the present body of knowledge. A preliminary examination of the literature suggested that a problem in aviation has always been accurately validating reliable measurement constructs related to training and then linking the constructs to the resulting flight deck behaviour (Parasuraman & Riley, 1997).

1.4 PURPOSE OF THE RESEARCH

The argument that pilots learn to fly aircraft in different ways is not new, and has been raised in the literature for some time (Sherman, 1997; Singh *et al*., 2005; Telfer & Moore, 1997). According to Vermeulen (2009), flight instructors are acutely aware of the fact that some students are deeply motivated to understand their aircraft, while others learn only the bare minimum so as to meet the minimum standards to pass the course. Similar conclusions were reached in a study of general aviation pilots (Telfer & Moore, 1997). Hence, some trainee pilots may be motivated by their instructors, together with trait factors such as a passion for aviation; whilst other students may see training merely as a means to an end. The way airline pilots learn to fly is a combination of their natural love for the activity and the subject of aviation on the one hand, and the learning environment on the other (Sherman, 1997; Tracey & Tews, 2005; Vermeulen, 2009). Only after quantitative measurement of the training environment as perceived by students, can the variables and phenomena responsible for individual behaviour on board the flight deck be thoroughly investigated. Similarly, pilots' attitudes towards training can also be competently explored only if their perceptions can be measured quantitatively (Naidoo, 2008).

In view of the aforementioned discussion, the main focus of the research was to develop an appropriate instrument to conduct a cross-sectional assessment of airline pilots' perceptions of the climate associated with advanced automated aircraft training at their organisation (airline).

Based on a comprehensive initial literature review, a hypothetical construct was developed which defined perceptions of the advanced aircraft training climate. In the context of the present study, and to meet the research objectives, a training climate was regarded as the prevailing conditions of the person (pilot), the group (instruction) and the organisation (airline) as experienced by the pilot during or after training to operate an advanced automated aircraft in the South African commercial aviation industry.

The primary purpose of the study was *to develop a valid and reliable instrument to measure airline pilots' perceptions of the training climate associated with advanced automated aircraft and to explore related phenomena statistically.* Overall, therefore, the fundamental goal of the study was to operationalise an unobserved hypothetical construct by developing a survey questionnaire (perceptions of the advanced automated aircraft training climate) based on three hypothetical latent sub-constructs (the person, the group and the organisation) that conceptualised the primary construct. The major organisational behaviour literature suggests that constructs are systemically correlated at three fundamental levels, namely at the micro, meso and macro levels of the organisation (Bott & Svyantek, 2004; Desler, 2002; Drucker, 1946). Furthermore, for the initial development of the hypothetical construct, this research study relied substantially on the premises sourced from prior psychology theory and organisational behaviour theory.

In addition, the purpose of the research was to analyse the perceptions of South African airline pilots as a specific unit of investigation, as people qualified to operate advanced technology aircraft for their companies. To fulfil the primary purpose of the study, a number of characteristics about the unit of analysis were then assumed. First, it was assumed that, as airline pilots acquire an increasing amount of industry experience, they move from being a dependent personality to being a self-directing entity. Second, it was decided that the units of analysis would all be current, qualified and fully trained airline pilots. Third, it was assumed that pilots use accumulated industry experience as a resource for learning. In other words, the sample under investigation use experience gained on other aircraft as a basis for learning new technology principles. Fourth, it was assumed that airline pilots' readiness to learn was oriented to the developmental tasks of their flight deck position.

As a secondary purpose, this study sought to explore the underlying structure of the relationship between the variables and attempted to explain the dynamic of the phenomena related to respondents' perceptions of the latent constructs, based on an appropriate level of statistical analysis.

1.5 RATIONALE FOR THE RESEARCH PROJECT

A preliminary analysis of the relevant literature suggested that there was no construct-validated measure of a training climate which was associated with advanced commercial aircraft and which was anchored in an organisational behaviour paradigm, even though it has been claimed that many important constructs in the behavioural sciences are organisation-based (Robbins, Odendaal & Roodt, 2004). Research studies that undertake to analyse the organisational learning environment can only be of value if cognisance is taken of the fact that the training climate is a critical component of resultant behaviour and therefore has a subsequent impact on overall organisational effectiveness. Using well-developed scales for the measurement of individuals' perceptions of a particular climate can determine human behaviour within an organisational system and context (Meister, 1999). Early seminal studies point out that individual human behaviour is latent within an enterprise and can significantly influence the overt (observable) aspects of an organisation at a very fundamental level (Likert, 1958). Applying such a theory in the context of the aviation industry could therefore benefit the current body of understanding.

A review of the literature suggests that future research is needed to explore the dyadic nature of a human-technology system in the airline industry. Meister (1999) argues that the measurement of aviation human factors should not be very different from the measurement of any other systems in the psychology environment. Indeed, some authors have suggested that specialised aviation metrics research would have real potential to contribute significantly to improving not only business, but also aviation safety (Funk & Lyall, 2000). One rationale for attempting this study was that analysis of the general understanding of the topic of interest over the past two decades tends to be grounded on fairly little positivist and quantitative research, so that there are vast gaps in what is currently known in both the organisational and psychological literature about climate constructs related to training for advanced aircraft (Funk & Lyall, 2000; Mitchell *et al.*, 2009; Parasuraman & Byrne, 2002; Sherman, 1997; Singh *et al.*, 2005; Wiener, 1993). For example, one study in this field examined perceptions of glass cockpits over a given period, and found that training aspects affected a significant portion of this perception. However, the authors of the paper acknowledged that the factors involved in airline pilots' perceptions of learning and training were a relatively unknown and new area of exploration (Mitchell *et al.*, 2009)

The aim of this research was to explore a relatively new area, marrying the principles of organisational behaviour, aviation training, human factors and advanced technology. This implies a more complex study, making an exploration of relatively unknown phenomena particularly difficult. However, the potential contribution to a scientific understanding of aviation psychology that could accrue from such an undertaking is so valuable that it vindicated the research approach. Moreover, the study investigated and determined the nature and content of a newly developed and empirically derived construct for the airline industry. The clarification of airline pilots' perceptions of the advanced automated aircraft training climate appears unique, because a review of prior research on the advanced automated aircraft training climate revealed little information on any training or learning measurement constructs. Similar constructs are, however, available in the general corporate training arena (Tracey & Tews, 2005). These were reviewed when conceptualising the present research study.

Overall, all searches on the current topic of interest revealed a general dearth of substantive aviation industry data. The fundamental importance of and rationale for this research aimed at airline organisations interested in aviation psychology is that new knowledge of the training climate can be used to develop adequate interventions. Such interventions include altering, modifying or enhancing current training methodology and paradigms, thereby improving organisational effectiveness, efficiency, competence and overall corporate competitiveness.

1.6 SCOPE OF THE STUDY

1.6.1 General scope

A review of relevant journals showed that research into aviation automation training had not yet produced the extensive empirical data necessary to bring about significant policy and regulatory change (Parasuraman & Byrne, 2002; Pohlman & Fletcher, 1999). Only one prior study has touched on advanced technology pilot training, namely that by Sherman, Helmreich and Hines (1995). However, these authors examined automation training only by specifically focusing on pilots' experiences with automated aircraft, using a previously developed framework by Wiener (1998). To date, the theoretical body of knowledge dealing with aircraft automation has been limited and has focused only on pilots' attitudes toward aircraft automation in general (Funk & Lyall, 2000; James *et al*., 1991; Naidoo, 2008; Sherman *et al.,* 1995; Risukhin, 2001; Wiener, 1998) and not on aspects related to training.

It was found that very little new information is available, particularly in the last five years. In fact, the bulk of the psychology and behavioural literature available on flight deck automation research spans only the last decade. Furthermore, no study has developed any specific instrument to critically assess or measure constructs associated with the training climate related to advanced automated aircraft grounded in organisational behaviour theory. The questions of *what* strategies airline pilots adopt and *why* airline pilots adopt certain pedagogical motives and strategies with regard to advanced automated aircraft have not really been asked thus far. The complex interplay of airline pilots' intentions in learning, the approaches they may choose, and the outcomes of their training, has not been empirically examined, creating a theoretical gap in aviation human factors research. The testing of new methods and procedures that measure perceptions accurately should be undertaken systemically – an argument championed by many renowned psychology scholars

(Cattell, Boyle & Chant, 2002). It has been emphasised in the literature that doctoral research in this specific area is appropriate and very necessary; so as to achieve the required levels of depth needed to make organisational changes (Sherman, *et al.*, 1995). Although the literature shows that the training of airline pilots contributes significantly to their perceptions of new technology aircraft (James *et al.,* 1991; Johnston, Fuller & McDonald, 1995; Naidoo, 2008; Wiener, 1998)*,* the measurement of perceptions relating to training and learning has not, to date, formed the basis for any study in this field*.* The scope of the current research was therefore based on the premise that an instrument should be developed empirically, starting from a sound theoretical foundation. However, it is a limitation on the scope of this research that such an approach has not been used in other research of this nature before, so that there is nothing to compare it to, resulting in highly exploratory methodologies.

1.6.2 Theoretical scope

How people see their learning environment is a fundamental antecedent to influencing the knowledge-gaining opportunities in, and effectiveness of, the training process (Schaap, 2000). Biggs's (1987) 3-P model of learning suggests that there are distinct issues regarding the strategic selection of either a deep, surface or achieving method of acquiring relevant knowledge for pilots learning to fly an aircraft. However, how *advanced automated* aircraft airline pilots perceive their learning environment has not yet been scrutinized using this framework.

Understanding general behaviour in a highly computerised flight deck has not necessarily improved safety in these aircraft (Mouloua, Gilson & Koonce, 1997). Furthermore, Mouloua *et al.* (1997) postulate that a pilot's perception of his or her comprehending advanced aircraft systems may be highly correlated with situational factors such as the subject's experience and the level of automation employed in the aircraft. However, a focus on specific behavioural traits, such as training or learning, should contribute to the body of knowledge and could have a significant impact on flight safety. An examination of perceptions of the training climate as experienced by the operators and dividing the research sample into specific demographic categorisations would result in the level of specificity required. Hence, the current theoretical scope of the study attempted to address more detailed issues by means of a thorough exploration.

The exploratory method used in this study relied on prior empirically derived theory. For example, based on Schaap's (2000) analysis of adult learning approaches, it was deemed important to ensure that the hypothetical measurement construct included concepts related to individuals' abilities to examine and reflect critically on their own learning process, because a quantitative link has been found between learning reflection and continued or sustainable effectiveness in a knowledge economy. To understand the effectiveness of airline pilots in their organisations, this method of theoretical reasoning was extrapolated. For instance, the theory may suggest that pilots who are not cognitively reflective of a higher order climate may not be capable of acquiring the relevant information to fly safely. This notion was combined with evidence presented by Telfer and Moore (1997), who reported that a lack of critical aviation-based knowledge significantly increased the probability that an individual would be the root cause of a serious incident or accident.

Correct reflection on learning processes will invariably promote an airline pilot's ability to handle new and complex situations. *Reflection* in this context and in terms of the theoretical scope used in the current study is then defined as the ability to integrate past learning experiences and apply this knowledge to new problems – an ability that has been shown to indicate a higher level of mastery in an acquired skill. Schaap (2000:xxvi) explains that the "way an individual views the process of learning influences the learner's approach to a learning opportunity and the effectiveness of the learning process concerned".

Some important psychoanalytical concepts are often neglected in examining organisational training. The scope of the current study therefore also included many principles taken from the fields of psychology and the behavioural sciences. For example, Aronson (1991) raises the following critical issues:

- the relationship between learning (training) and the stages of development (in this case, the levels of experience of airline pilots), which could emphasise one or more modes of learning;

- the role of values, goals and ideals in learning strategies;
- the relationship between personality, the self and the organisation in the learning process; and
- the psychodynamics of thinking, in other words, motivation theories and both primary and secondary processes which contribute to theory building in the fields of thinking and learning.

The prevailing training climate (trainees' intentions, approaches and outcomes) as perceived by advanced automated aircraft pilots was explored as a human factor issue in the current study, following on from suggestions derived from important field research by James *et al*. (1991) and Naidoo (2008). It is necessary to measure perceptions in order to reduce the human factors knowledge gap in aviation psychology and also to enhance academic understanding of the related phenomena, which are significant for flight safety.

Inspired by Schaap's (2000) model, which showed distinct systemic linkages between learning approaches, learning environmental factors, personal factors and learning outcomes, the theoretical scope of the research was expanded to include systemic models. According to Drucker (1946), an organisation is a social entity and a repository of knowledge; furthermore, one must regard this entity as interdisciplinary and as displaying some philosophical sophistication. The mechanisms operating between aviation-related learning components and systemic organisational behaviour were examined further in the literature study. This approach guided an understanding of the psyche of an airline pilot (who has been trained to fly an advanced automated aircraft) and the development of a fundamental psychological measurement instrument. The phenomenological components that were identified as encapsulating the knowledge environment of a modern airline pilot were derived from this theoretical scope.

1.7 RESEARCH OBJECTIVES

The research problem and objectives of the study were finalised only after a comprehensive and non-exhaustive preliminary literature review, which revealed a need for the development of an assessment instrument containing the dimensions of the particular construct of *perceptions of the advanced automated aircraft training climate*.

The primary objective of the study was therefore to obtain an empirical estimate of the hypothesised construct by constructing a multi-dimensional questionnaire in order to develop a valid and reliable measurement scale. The following research objectives were generated to guide the study:

- to identify from the literature which organisational behaviour attributes apply to the main and sub-constructs;
- to develop a hypothetical multivariate psychological systems model (founded in empirically grounded theory) from which criteria for the construct of *perceptions of the advanced automated aircraft training climate* could be identified and tested in an quantitative study involving a cross-organisational sample of airline pilots from South Africa;
- to generate a tentative pool of scale items based on a model of the construct;
- to validate the items of the hypothesised construct statistically, by quantitatively examining the judgements gained from subject matter experts and using Lawshe's (1975) content validity ratio technique;
- to obtain sufficient empirical data to explore the nature of the latent factors of the main research construct and to develop an understanding of their relationships to the surface attributes and to each other;
- to statistically develop a valid and reliable measurement scale based on the main research construct; and
- to explore the statistical relationship between respondent variables and the latent factors of the construct.

1.8 OUTLINE OF THE STUDY

The study is organised as follows:

Chapter 1 has introduced the study by demonstrating the current gap in the knowledge in the aviation industry and providing an orientation regarding the research purpose, scope, objectives and focus of the research.

Chapter 2 sets out a comprehensive literature review in terms of the human-machine interface. Firstly, important terms are defined. Secondly, in order to maintain a relatively logical flow, the literature review was structured around the following key themes: advanced aircraft technology and its impact on human behaviour and *vice-versa*, learning measurement and the theoretical approaches adopted by trainees, examining the training climate construct based on organisational behaviour and relevant psychology theory; which then leads into a discussion of the advanced automated aircraft climate and relevant specific training aspects for this unique work environment. The literature on these themes was synthesised with the aim of providing an integrated theoretical understanding and critique of the topic.

The aforementioned review is comprehensive, but far from exhaustive – only pertinent and relevant theories (arguments) were selected for this critical evaluation of the current body of knowledge. The literature review supports the conceptual development of the main hypothetical construct, and thus forms the core foundational phase of this research.

Chapter 3 links the literature from Chapter 2 in terms of advanced aircraft technology and training. In addition, this chapter critiques the work from the existing body of knowledge to underpin and contextualise the development of a tool to measure perceptions of the advanced aircraft training climate. The chapter also states the delimitations of the present study and seeks to clarify the research construct for statistical analyses in Chapter 4 and Chapter 5.

Chapter 4 describes the research and statistical methodology followed, which includes the steps taken in the final scale design and development. A discussion of the logic of the sampling technique, surveying method, questionnaire construction, factor analysis, reliability and item analysis forms a major part of this chapter. Furthermore, Chapter 4 discusses the initial statistical examination of the main research construct. The results of the Lawshe (1975) technique is discussed and presented after thoroughly examining judgements from subject matter experts. Therefore, Chapter 4 discusses and finalises the operationalization of the main hypothesised research construct.

Methodological depth is achieved within this section, because the chapter also attempts to defend and substantiate the choices made regarding the use of parametric and non-parametric statistics from a theoretical perspective. In addition, effect sizes and practical significance, as relevant to Chapter 5 are summarily discussed.

Chapter 5 explores the results of the study by examining the outcomes of the item, factor analyses and logistic regression analysis. The chapter sets out in detail the analyses and results on the total data set by examining the effects between selected demographic variables and the latent factors of perceptions of the advanced aircraft training climate construct, by using descriptive, comparative and associational statistical methods.

Chapter 6 revisits the research objectives, methodology and limitations of the study in light of the final results. The main components of the research construct are summarised and the managerial impact of the demographic variables on the outcome factors are reviewed. The chapter also proposes recommendations for future research.

CHAPTER TWO:

LITERATURE STUDY – THE HUMAN-MACHINE INTERFACE

2.1 INTRODUCTION

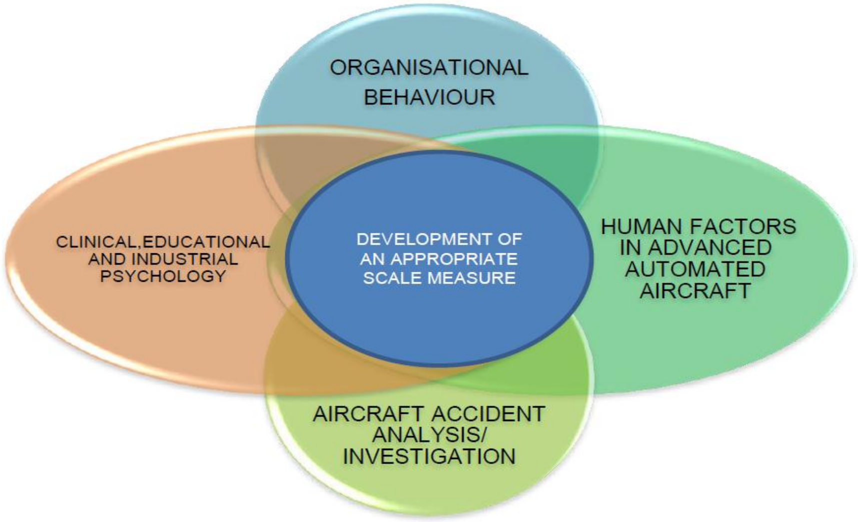
When a study is exploratory in nature, it is crucial to conduct a thorough literature review (Mouton & Marais, 1996). Reviewing the appropriate literature and examining critically, related prior research, can provide a good indication of where the current thesis fits into the context of the present body of knowledge. According to Babbie (2010), an effective review of the literature consists of evaluating selected documents on a given research topic. Human factor research in aviation is a relatively neglected topic when compared to other areas in psychology and organisational behaviour (Dekker & Johansson, 2000). For this reason, much of the literature consulted is at times, as much as two to three decades old. It is therefore intended that the present research study would add new material to the current knowledge deficit, with useful information.

In this chapter, the nature of the interaction between human beings and highly advanced technology in an aviation industry setting is examined; more specifically, the impact of advanced aircraft on human behaviour is discussed. A multi-pronged approach was followed in analysing the evolution of advanced aircraft pilots' current working environment.

Four decades ago, Gordon Moore postulated that the number of transistors on a silicon chip would double every two years (Voller & Agel-Porte, 2002:699) – he claimed that "[a]nother decade is probably straightforward...there is certainly no end to creativity" (Moore, 2003). This prediction has remained true thus far – exponential advances in technology were made possible and the benefits from miniaturised components still continue to proliferate everyday life. The cumulative impact of such advances has unquestionably revolutionised the electronic world. More importantly, especially for the purposes of the current study, it has drastically changed the face of the aviation industry. Commercially available passenger and cargo aircraft employing highly complex computerised automation has become much easier to engineer and manufacture (Australian Transport Safety Bureau, 2007). As in many industries, new technology has always had a significant impact on the lives of those who earn a living from working with it – in the case of the aviation industry, on the attitudes, skills and proficiencies of pilots (Abbott, 2010). As with many technologically advanced products found in today's marketplace, some individuals can easily embrace the necessary learning of the skills required to operate the new product effectively, while others find it less easy to adapt (Naidoo, 2008).

In order to understand the context in which the development training measures are used, it is essential to study the level of aircraft automation. This will enable comment and increase the applicability of the results and recommendations of the study. In this chapter, the advanced and highly automated aircraft is introduced. Thereafter, the chapter critiques the human factor element associated with the technology in order to provide a background for the design of the hypothetical measurement construct intended to meet the research objectives. Figure 1 depicts a contextual framework for the present study, by graphically synthesising the areas covered in the literature review.

Figure 1: Synthesis of the literature study



2.2 CONTEXTUAL DEFINITIONS

Definitions of two key terms used throughout the research are provided in Table 1. The contextual definitions aid in the discussion and literature review which follows and make it easier to grasp the meaning of fundamental aviation automation concepts.

Table 1: Definitions of some key terms

Term	Definition
Advanced Automated Aircraft	According to Risukhin (2001), the advanced automated aspects of aircraft consist of two main components, namely: • the computerised flight deck systems, for example, the flight director/autopilot and the flight management system; and • the computerised airframe and mechanical subsystems, for example, the electronic engine control, propulsion, auto-throttle and auto-thrust functions.
Advanced flight deck or glass cockpit	The glass cockpit is "a system of cathode ray tubes of liquid crystal display flat panels that provide key critical information and control through advanced computers about the status of the aircraft" (Wiener, 1988:10).

2.3 CHARACTERISTICS OF ADVANCED AUTOMATED AIRCRAFT

According to Airbus (2011b:1.22.10), the general philosophy that underpins automating an aircraft is that doing so "reduces cockpit workload, improves efficiency and eliminates many routine operations normally performed by the pilots" in the normal flight envelope. Various scholars (Parasuraman & Byrne, 2002; Sarter & Woods, 1994; Sherman, 1997; Wiener, 1988) provide a similar explanation in terms of defining the automation of aircraft flight decks. In addition, the optimum use of aircraft automation involves the integrated and co-ordinated manipulation of the following basic aircraft components:

- the autopilot;
- the flight director;
- the auto-throttle or auto-thrust systems; and
- the flight management system.

The most advanced aircraft today offer users a fully automated system in terms of both the lateral and vertical profiles (Airbus, 2011a). Ascending levels of computer-based automation provide the flight crew with an ever-increasing number of options and strategies to choose from. The choice of automation options is complex, because it must be accomplished in accordance with the particular task at hand. For instance, tactically complying with air traffic control requirements in the short-term when in close proximity to the airfield, versus, strategically programming the flight management system for long-term navigational requirements so as to safely and efficiently traverse a continent (Parasuraman & Byrne, 2002).

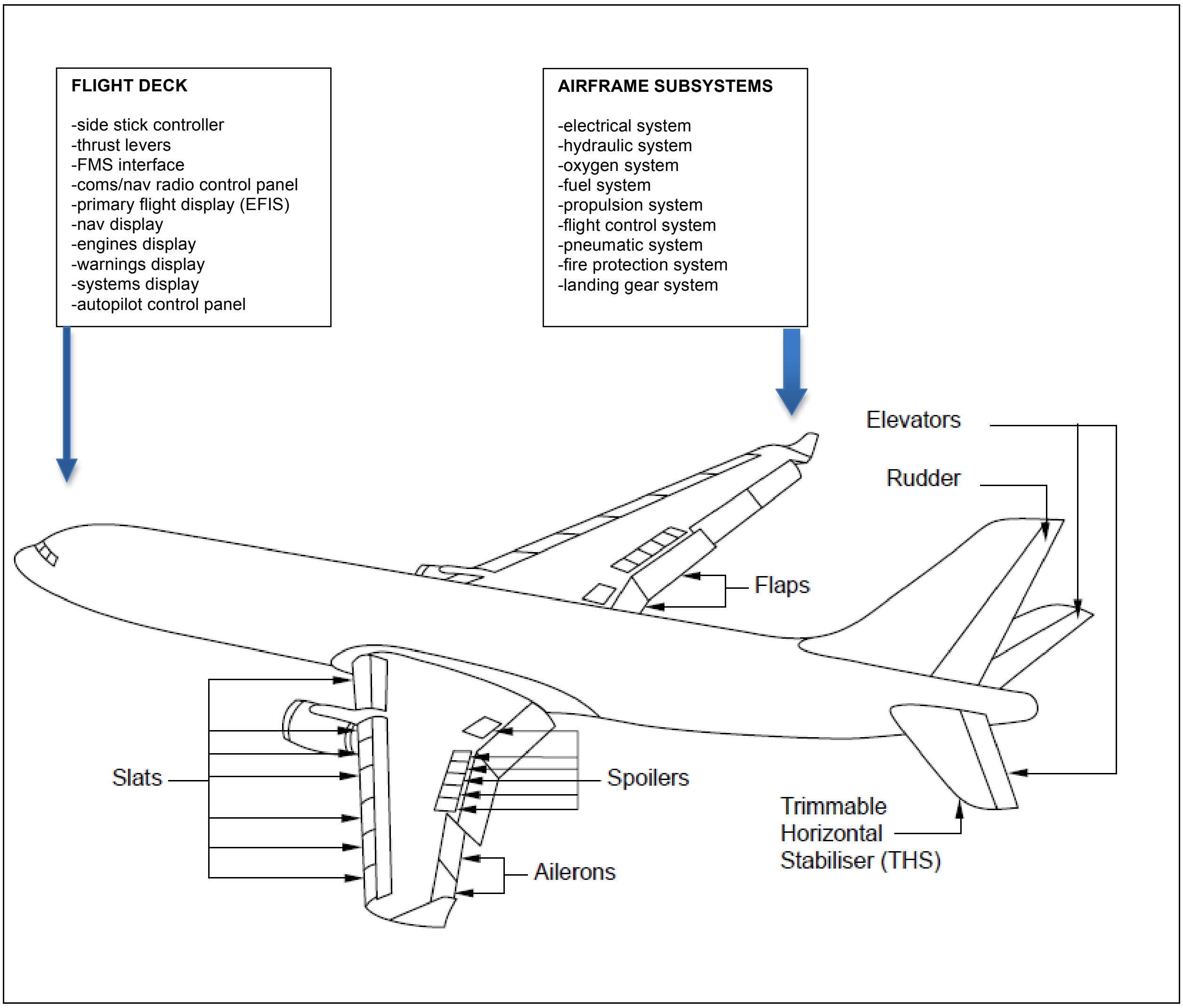
An advanced automated aircraft contains a multitude of primary systems and peripheral subsystems. Flight deck crewmembers that learn to fly such modern aircraft are required by law to understand and comprehend the working details of the aircraft in totality. This includes the on-board computerised flight deck systems and other advanced airframe-related subsystems (South African Civil Aviation Authority, 2011).

When the level of knowledge required by an airline pilot is considered, it is clear that the measurement of any hypothesised construct of airline pilots' perceptions of the climate associated with advanced automated aircraft training at an organisation (airline) is multifaceted and complex. For this study, it was determined that any hypothesising of relevant constructs should encompass an assessment of pilots' perceptions of the climate associated with an entire transition-training course. A full transition-training course consists of understanding two main components of an advanced automated aircraft, namely

- the flight deck systems; and
- the airframe mechanical subsystems.

These two components of the modern aircraft are depicted in Figure 2.

Figure 2: The two main components of an advanced automated aircraft



Source: Adapted from Airbus (2011a) and Risukhin (2001)

For the purposes of this study, advanced automated aircraft refer to the two main components of the machine (computerised flight deck systems and computer-based airframe systems) in combination.

Figure 2 is a model of an Airbus A330 variant, one of the world's most advanced commercial jet aircraft. Some of the fundamental systems of each main aircraft component are shown in the figure, illustrating the complexity of the components. Each subcomponent of the aircraft system requires specific levels of understanding by pilots, so that routine and non-normal operations can be performed safely. A lack of comprehension, or significant knowledge gaps regarding some technical aspects of the aircraft, can result in an unsafe outcome (Ribbens & Mansour, 2003). It is important to note that the evolution of medium to large commercial aircraft over the last fifty years has resulted in pilots' requiring an increased ability to grasp specialised technical concepts. The most important of these are discussed in more detail within the next section.

2.3.1 Computerisation of aircraft systems

Rapid improvements in modern digital electronics have resulted in equally advanced improvements in aircraft design and implementation. Tooley (2007:1) contends that a "modern aircraft simply could not fly without the electronic systems that provide the crew with a means of controlling the aircraft". A technologically advanced commercial jet aircraft is dependent on numerous computer-based systems for an exceedingly broad variety of operations, from flight control and instrumentation, to navigation, communication and electronic engine control. On-board computers sense and indicate the aircraft's trajectory in relation to the earth, the aircraft's heading, altitude and speed. Thousands of invisible sensors in and around the aircraft work in complete harmony with both microscopic and large mechanical devices. Sophisticated computers process a plethora of environmental and internal information, ultimately contributing to a complex and almost "living" entity, which may be, why sometimes, an aircraft appears to have a "mind of its own" (as one survey respondent to this study commented). Without the digital logic provided by highly advanced microprocessors, these types of aircraft would not leave the ground (Tooley, 2007).

Resource limitations and increasing corporate competiveness have ensured that aircraft manufacturers research and build aircraft that take advantage of the benefits of modern micro-processing power. However, the potential for the limitless incorporation of new and sophisticated engineered components in aircraft has raised a number of human-centred concerns amongst experts in the field (Barker, 2011). Critics are concerned with the rate at which technology is being incorporated into the modern flight deck, with the result that there is increased detachment in the human-machine interface (Abbott, 1995; Barker, 2011). Observers have gone so far as to label increased reliance and perhaps over-reliance on aircraft automation and computerisation an *addiction* (Barker, 2011). The current lack of understanding between technology and human behaviour has caused a widening gulf between the opinions of two significant groups of people, namely airline pilots and engineers (Poprawa, 2011). Engineers seek to reduce the need for any human intervention during aircraft operations, whereas pilots seek to gain more control over and flexibility of the aircraft (Sarter & Woods, 1994). The level of computerisation in the modern aircraft is likely to present many challenges and raise debate in both the technical and psychology fields in the foreseeable future. The debate at the future human-human level in aviation, as opposed to the human-machine level of interaction, may prove to be an interesting research topic for further research, however, falls beyond the scope of the current topic.

2.3.2 The dominance of aircraft technology

An analysis of the literature shows that there is some consistency in the various attempts to formulate what constitutes advanced flight deck automation. In most instances, authors share the notion that flight deck automation involves a gradual handing-over of power from the human operator to the computer system (Parasuraman & Byrne 2002; Sarter & Woods, 1994; Sherman, 1997; Wiener, 1988). Increased use of computer processing power in aircraft has given rise to the term "glass cockpits" (Taylor & Emanuel, 2000:18). Table 2 compares the various convergent definitions found in the literature. A number of concerns have been raised by human behaviour experts, regarding pilots' control and management of advanced aircraft and the transition required to adapt to new aircraft technology (Barker, 2011; Lyall & Funk, 1998; Sarter & Woods, 1994). By altering the roles of the operator and the machine, advanced automation has increased efficiency, whilst simultaneously extending the available human capability (Schutte, 1998). Much empirical evidence provided by scholars investigating this area began contradicting the utopian promise that increasing flight deck automation increases safety linearly (Parasuraman & Riley, 1997; Rigner & Dekker, 2000; Sarter, 1996). The new responsibilities left to the human operator have resulted in new mistakes, errors or omissions (Lowy, 2009). A paradoxical decrease in situational awareness, increased mental workload, poorer efficiency in systems monitoring and a degraded ability to intervene during an automation failure are some of the concerns that have been cited. Interestingly, all of the major concerns being raised in the literature with regards to advanced aircraft, link the human being directly. The fact that there may be a distinct lack of understanding or comprehension of critical (more complex) technical topics by advanced automated aircraft pilots is a reason for the past and present concerns (Lyall & Funk, 1998; Poprawa, 2011). Some studies have noted a negative correlation between factors such as understanding and comprehension, and pilots' perceptions of advanced flight deck automation (Naidoo, 2008). In other words, the increasing dominance of complex systems may result in a reduced ability to understand these aircraft. Problems with understanding technology or pilots' lack of actually comprehending aircraft system complexity can be linked to the effectiveness or ineffectiveness of organisational training efforts (Moore, Po, Lehrer & Telfer, 2001). Researching the phenomena associated with such organisational efforts also provides a reason for further perception studies involving advanced aircraft.

Table 2: Definitions of advanced flight deck automation

Source	Definition of advanced flight deck automation
Wiener (1988:436)	Flight deck automation is "when some tasks or portions of tasks performed by the human crew can be assigned, by choice of the crew to machinery". Cockpit automation is also "regarded as computational support allowing some procedures to be omitted by the crew".
Sarter and Woods (1994:5)	Flight deck automation is "the allocation of functions to machines that would otherwise be allocated to humans".
Sherman (1997:2)	Flight deck automation is "the replacement of a human function, either manual or cognitive, with a machine function".
Parasuraman and Byrne (2002:315)	Flight deck automation is "the gradual and increasing replacement by machines and computers of functions once carried out by flight deck crew".

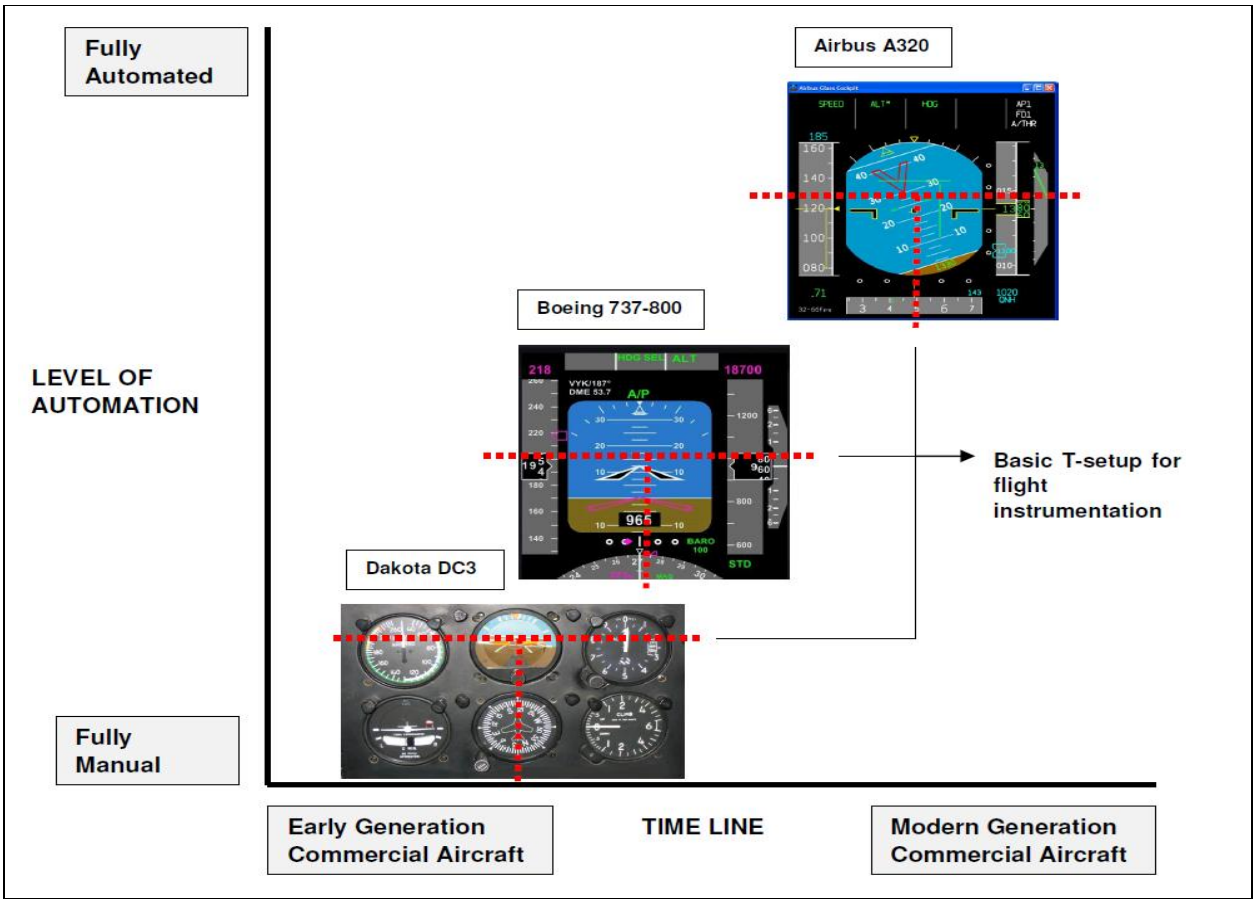
2.3.3 The advanced flight deck

Advances in technology have changed the appearance of flight decks substantially. A core difference between an analogue flight deck and a modern digital or glass flight deck is that in the glass set-up there is extensive use of electronically generated graphic displays (Airbus, 2011b). These displays are also coupled with underlying computer sensors, electronic circuitry and software. This remarkable evolution of the flight deck (from analogue instrumentation to digital instrumentation) is the most noticeable and tangible difference between older aircraft and the modern generation of aircraft (Risukhin, 2001). Because the most significant differences between the various periods of aircraft design can be found in the flight deck itself, most of the subsequent discussion relates to these differences. In addition, a significant proportion of automation issues and problems in the human-machine interface occur in relation to flight deck systems rather than in relation to airframe sub-systems (Parasuraman & Byrne, 2002).

The digitised flight deck system effectively and efficiently replaced the earlier analogue system in commercial aircraft approximately fifteen years ago (Ribbens & Mansour, 2003). The quantum leap in aircraft technology since the inception of flight has resulted in a series of revolutionary changes in the basic flight deck layout. The most obvious changes in aircraft design are noticed in the cockpit – or more correctly, the flight deck (when one refers to a commercial transport airliner). These are overt design changes, which give rise to the *glass* concept, whereas the peripheral mechanical subsystems of the advanced aircraft airframe may be considered latent changes in design. However, the latest generation of commercial jet aircraft, such as Boeing's 787 variants have seen a radical airframe update. These aircraft now boast noticeable wing and fuselage design changes, which are claimed to make the aircraft far more fuel-efficient and environmentally friendly (Boeing, 2009). Figure 3 depicts the advances, which occurred in flight deck instrumentation design during this transition (the timeline on the horizontal axis shows the independent variable, with the level of automation on the vertical axis as the dependent variable).

In order to maintain the correct and efficient scan of primary flight instruments, the initial design of the instrument display in an aircraft features a T format (Abbott, 2010). In this format, critical indicators such as the aircraft's lateral speed, trajectory and vertical speed are found at the top, and directional indicators such as the compass and turn-and-slip indicator are positioned below the T-bar. Although the basic T set-up of primary flight instruments has remained unchanged for some time, Figure 3 clearly depicts the significant ergonomic and aesthetic evolution in the primary flight instrumentation of the glass flight deck. The use of digital displays provides pilots with a tighter clustering of important flight information, resulting in improved situational awareness. Engineers are likely to continue to use the T set-up for displaying flight information to pilots, because "maintaining the relevant positions of the instruments has been important in allowing pilots to adapt from one aircraft type to another" (Tooley, 2007:11).

Figure 3: Evolution in primary flight instrumentation



Source: Adapted from Lyall and Funk (1998), Ribbens and Mansour (2003) and Tooley (2007)

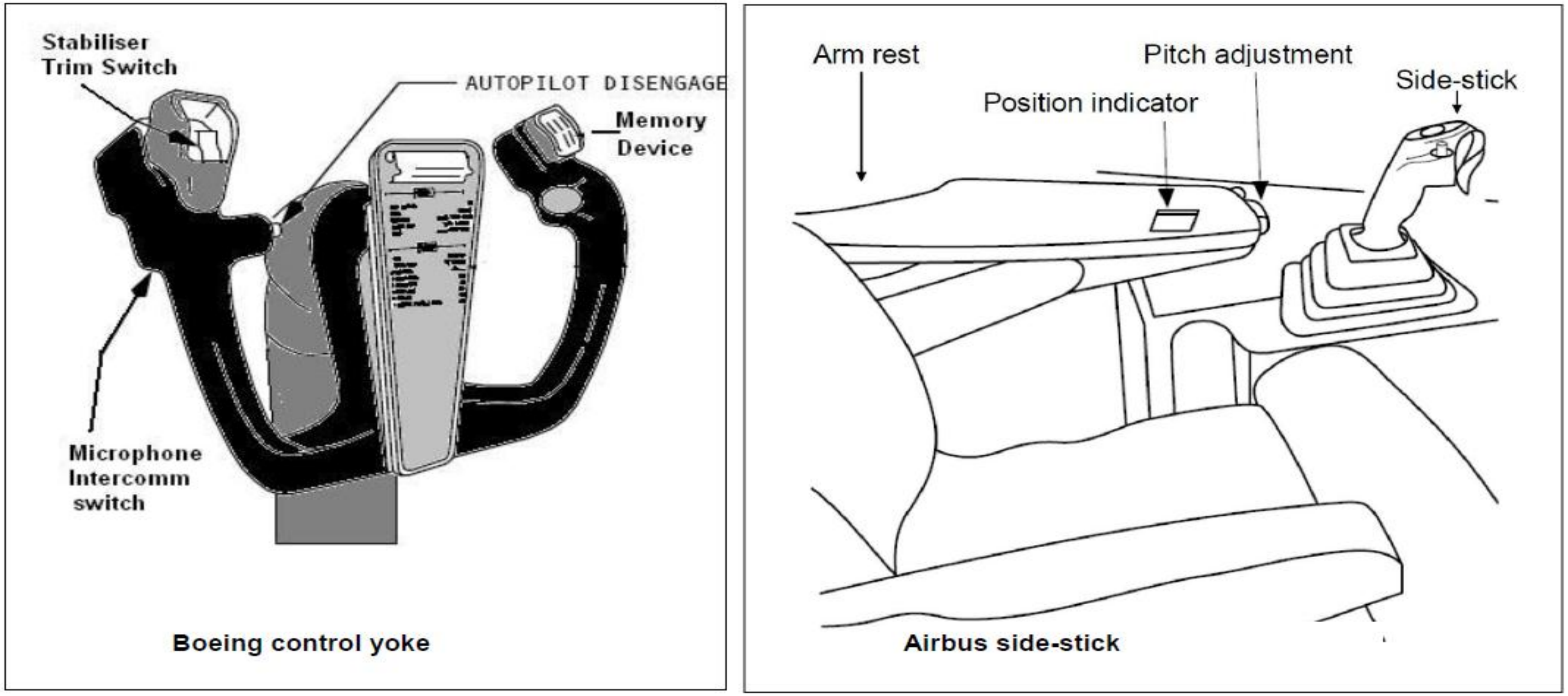
Other design changes in the modern flight deck have occurred in the way pilots manipulate the aircraft flight control surfaces. A critical difference between the design philosophies adopted by the two largest commercial aircraft manufacturers, Airbus and Boeing; is that Boeing continues to use a central control column as a means of manual flight control (Figure 4). By contrast, present-day Airbus commercial jet aircraft use an extremely sophisticated piece of technology, namely the side-stick, for manual control of the aircraft (Airbus, 2011a). The logic incorporated in the side-stick is highly complex, but it has proved to be an invaluable tool that has improved both ergonomic efficiency and aircraft safety characteristics. Nonetheless, experts in the field tend to disagree on which of the two systems is safer or more effective (Barker, 2011; Bent, 1996; Helmreich, 1987).

In normal flight, Boeing provides the pilot with comparatively more manual intervention from the central control column, while the Airbus side-stick system remains semi-automatic (Figure 4), and never allows for full manual control by the pilot (full manual control can only be achieved through the horizontal stabiliser for pitch, and rudder control for yaw, when in direct law, that is, after the failure of all the flight control computers, which is highly unlikely).

Boeing continues their philosophy of retaining a central control yoke in their most advanced aircraft to date, the B787, although the central control column is now based completely on fly-by-wire technology (advanced computerised automation, therefore there is no mechanical link between the control yoke and the aircraft flight control surfaces. Also see Figure 7).

The Boeing manufacturer premise is that commonality is sustained between the company’s family of 777s and 787s, allowing for quicker conversion (transition) type training by maintaining the more conventional control column design. By contrast, Airbus Industries have argued that the removal of the central column means that a pilot’s view of the primary instrumentation is no longer restricted, thereby providing a superior ergonomic state (Hradecky, 2011).

Figure 4: Comparison of specific flight control mechanisms

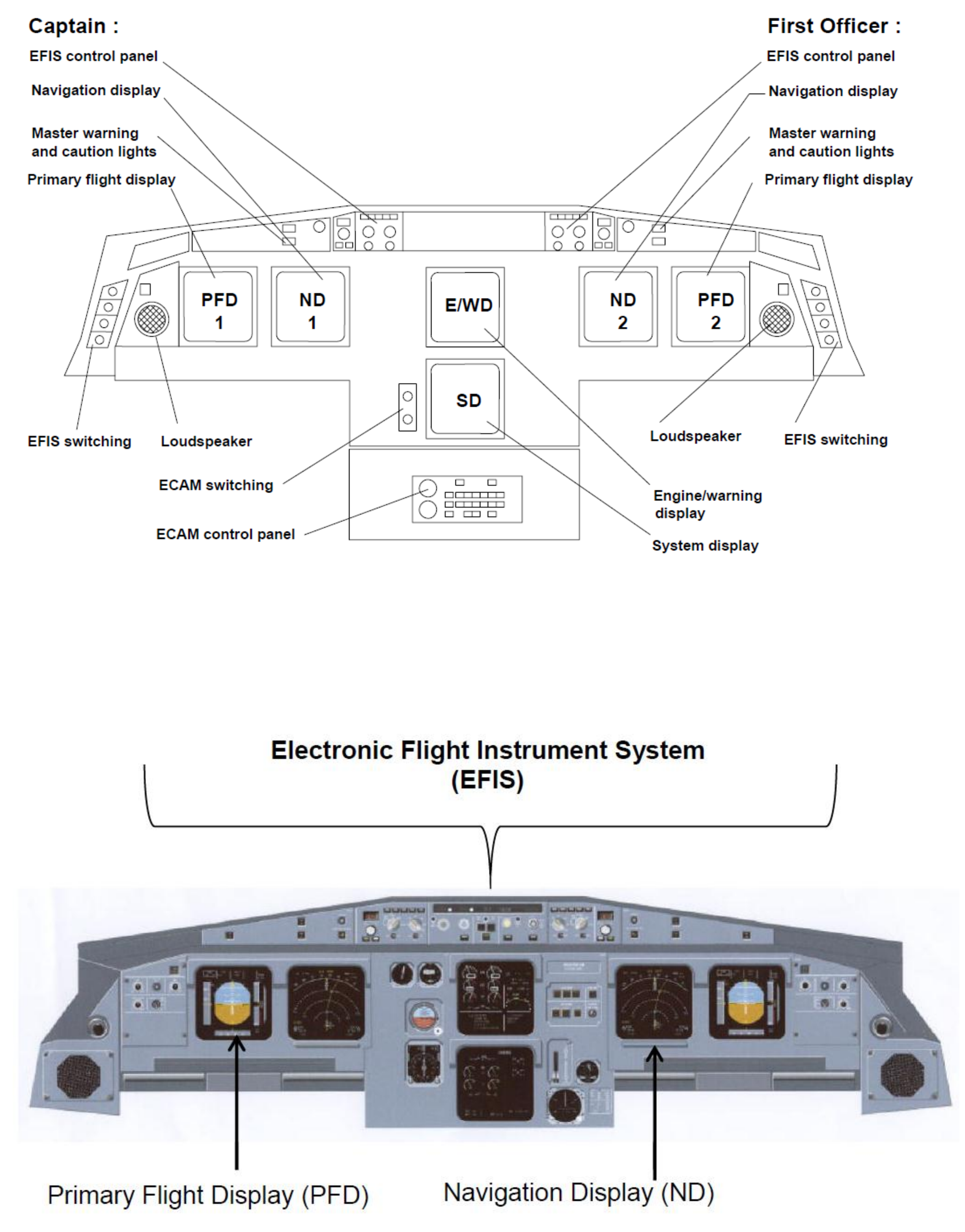


Boeing control yoke

Airbus side-stick

Sources: Adapted from Airbus (2011a) and Boeing (2010)

Figure 5: Advanced flight deck instrumentation console (Airbus A320 example)



Source: Adapted from Airbus (2011a)

Figure 5 illustrates the advanced flight deck instrumentation console housing the *glass* display system. In an advanced aircraft such as the Airbus A320, important displays for aircraft control are integrated within the EFIS (Electronic Flight Information Systems), while engine parameters, cautions, warnings and emergency procedures are integrated within the ECAM (Electronic Centralised Aircraft Monitoring) system, which was once the domain of a separate crewmember, namely, the flight engineer. Such a set-up has made it possible to operate even the largest passenger commercial aircraft with only two crewmembers. Furthermore, the overall system is designed in such a manner that appropriate information is presented in a timely and arguably more effective manner to the pilots. The principle of the display design is that it prevents an overload of unnecessary incoming information, therefore making it possible to operate the aircraft with less crewmembers then ever before (Parasuraman & Byrne, 2002).

The advanced flight deck layout also illustrates how designers have kept the most important flight information displays in a familiar configuration (see also Figure 3), allowing pilots to adapt more easily when switching to different models (Parasuraman & Byrne, 2002). This set-up attempts to harness ergonomics in an effort to reduce human-factor related problems and to improve safety measures in a technologically advanced cockpit (Sarter & Woods, 1994).

In addition to the difference in cockpit or flight deck setup, the response of the aircraft's actual control surface deflections after a pilot input, is considered artificial in a modern fly-by-wire aircraft (Hradecky, 2011). According to Dole (1989), the basic control of an aircraft (be it conventional or fly-by-wire) is a product of aircraft pitch (rotation about the lateral axis) and roll (rotation about the longitudinal axis). Some authors then propose that in a conventional aircraft, the pilot has a *sense* of the aircraft from a direct *feel* of control surface deflection, whereas; in the more modern fly-by-wire aircraft this *feel* is artificially generated by computer algorithms to provide feedback to the pilot depending on the current phase of flight (Risukhin, 2001:81-83) (discussed further in section 2.3.4). A comparison of the conventional and the fly-by-wire aircraft (in this case an Airbus A320/A330/A340) handling characteristics are best tabulated in terms of pitch and roll (see Table 3).

Table 3: Conventional and fly-by-wire aircraft control comparison

Pitch	Conventional flight controls	Fly-by-wire flight controls
Pitch Rate	• Pitch rate will vary in terms of control surface displacement and airspeed. • Aircraft pitch is unaffected by loss of airspeed information.	• Pitch rate is commanded by the equivalent in G-loading. • Pitch rate is unaffected by loss of airspeed information.
Aircraft response	• Aircraft response differs at varying airspeed. • Response is unaffected by loss of airspeed information.	• Aircraft response is the same at all airspeeds. • Response is unaffected by the loss of airspeed information.
Aircraft trim	• Trim is manual, and becomes more sensitive with an increase in airspeed. • Trimming is unaffected by loss of airspeed information.	• Aircraft trim is completely automatic. • Trim is unaffected by the loss of airspeed information.
Control column feel	• An artificial feel is introduced to simulate increased stick force at high airspeeds to prevent pilot over-controlling. • Here the basic introduction of an artificial feel can be unrepresentative of actual aircraft speed.	• The control column has the same feel at all speeds. • The feel is unaffected by the loss of airspeed information.
Aircraft envelope protection	• No flight envelope protection. • Unaffected by the loss of airspeed information.	• Full flight envelope protection is provided. • With the loss of airspeed information, the computer can only provide a G-load demand protection.
Roll	Conventional flight controls	Fly-by-wire flight controls
Pitch Rate	• Roll rate will vary in terms of control surface displacement and airspeed. • Aircraft roll is unaffected by loss of airspeed information. However, control limiters may be affected.	• A roll rate is commanded by the pilot’s control stick. • Roll rate will vary in terms of airspeed and with control surface displacement, adjusted for aircraft configuration, when airspeed information is lost.
Aircraft response	• Aircraft response differs at varying airspeed. • Response is unaffected by loss of airspeed information.	• Aircraft response is the same at all airspeeds. • Response will vary with a loss of airspeed information depending on actual airspeed and aircraft configuration.
Control column feel	• Aircraft feel is the same at all speed, however, limiters would change the response at high speeds. • Feel is unaffected by the loss of speed information, except for the effect from control surface limiters.	• The control column has the same feel at all speeds. • The feel is unaffected by the loss of airspeed information.
Aircraft envelope protection	• No flight envelope protection. • Unaffected by the loss of airspeed information.	• Full flight envelope protection is provided. • With the loss of airspeed information, the computer cannot provide roll protection.

Source: Adapted from Poprawa (2011)

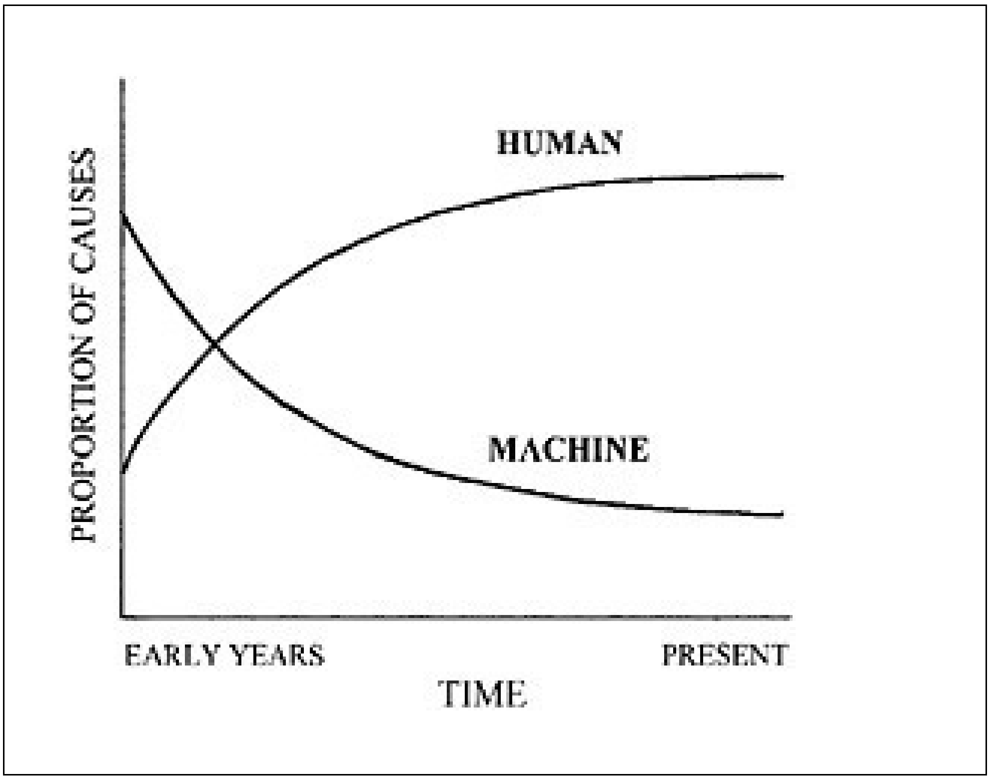
The tabulated comparison (see Table 3) illustrates that the pervasive nature of computerisation into aircraft flight control ensures that the workload associated with operating the modern fly-by-wire aircraft is far less than in the conventional aircraft (Poprawa, 2011). However, Barker (2011) challenges this notion, by arguing that the new ease in aircraft control and overall workload, may manifest in pilot complacency or overdependence on the protections provided for by the computer-based systems. This presents the argument that when computer dependent protections are lost, pilots may find themselves in unfamiliar territory and unable to control the aircraft safely (Cockburn, 2007; Hradecky, 2011).

A search of the currently available database reveals that more research is required to ascertain the level of a pilot’s actual situational awareness loss with a loss in computer-based system protections. Bent (1996) has proposed that only from superior advanced aircraft training can there be assurances that pilots will remain competent whenever there may be a loss of protection provided for by advanced automation. Nonetheless, available statistics have shown that high automation in aircraft coupled with superior ergonomic flight deck design has resulted in far safer and financially viable air travel (Boeing, 2009; 2010).

2.3.4 Advanced airframe and mechanical subsystems

In the evolution of modern advanced aircraft, changes to the flight deck are conspicuous and very impressive. However, similar advances in peripheral systems, which constitute the advanced aircraft, are often unseen and hence neglected (Ishida & Kanda, 1999). Pilots’ misunderstanding or deficient knowledge loops in terms of aircraft systems have led to fatal accidents in the recent past. For instance, confusion about the auto-thrust system of an Airbus A320 contributed significantly to a TAM air crash in São Paulo, Brazil (NTSB, 2009). Chambers and Nagel (1985), as well as Koonce (2003), suggest that the mechanical elements of aircraft have become extremely reliable and can only conclude that the majority of accidents and incidents are significantly related to avoidable negligent human behaviour (see Figure 6). In other words, fewer accidents or serious incidents may be attributed directly to the aircraft itself.

Figure 6: The relationship between mechanical failures and human factors



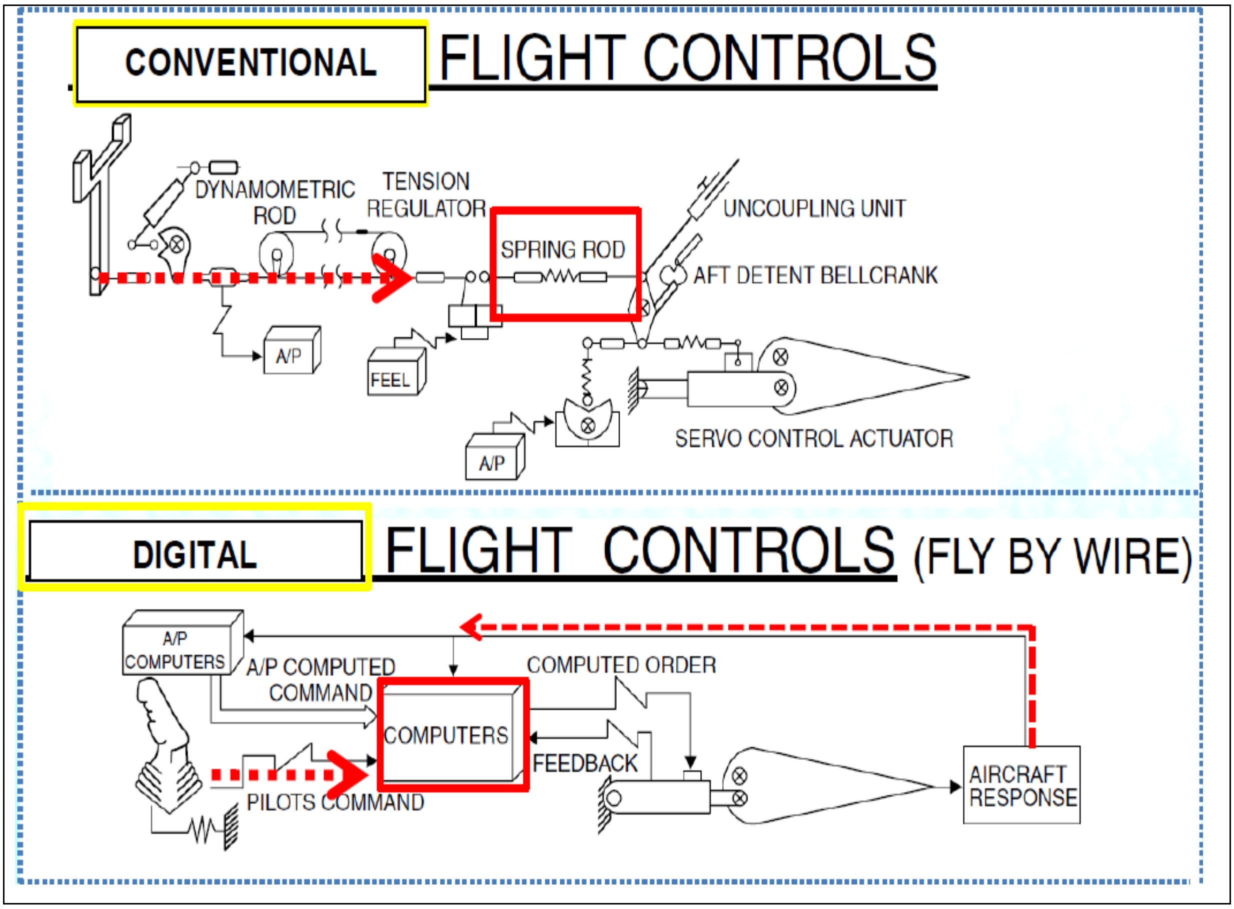
Source: Chambers & Nagel (1985)

The greatest changes to aircrafts' peripheral and airframe mechanical design arguably came in the form of the digital electrical flight control systems, also referred to as fly-by-wire (FBW). Briere and Traverse (1993) discuss how these computer-based, fault-tolerant (high redundancy) systems have enhanced the safety aspects of aircraft flight control substantially. Such a system first appeared in the Airbus A320 in 1988. Since then, more manufacturers have opted to computerise such mechanical subsystems, in the hope of increasing both safety and the savings resulting from greater efficiency. Pilots are now able to fly aircraft with unprecedented precision and accuracy, saving both time and money (Airbus, 2011b).

Yeh (1996) found that Boeing's triple seven (B777) aircraft's primary flight control system exhibited high levels of redundancy: "The heart of the FBW concept is the use of triple redundancy for all hardware resources: computing system, airplane electrical power, hydraulic power and communication path" (Yeh, 1996:294). However, this increase in efficiency, coupled with aviation safety requirements, derived from advanced computer programming, software and hardware, requires a paradigm shift in pilots' comprehension of systems. It is possible that changes can result in adverse human factor issues and have led to unforeseen problems according authors such as Billings (1997), Ishida and Kanda (1999), Parasuraman and Riley (1997) and Mitchell *et al*. (2009).

Figure 7 depicts the general differences in aircraft control design at a mechanical level. The actual mechanical linkages have now been replaced by a digital signal propagated through an electric wire in the modern aircraft. Furthermore, the figure depicts how manufacturers went from a direct link between the pilot and the flight control system to an indirect link, controlled and monitored by sophisticated computer-based hardware and software. One concern with these advances is that some accidents are now being attributed to incorrect pilot control of advanced FBW systems (Koonce, 2003; NTSB, 2009). The very design that was intended to prevent accidents is now being singled out as a major contributory factor to accidents.

Figure 7: Comparison of two aircraft control system types



Sources: Adapted from Airbus (2011a); Briere and Traverse (1993) and Yeh (1996)

2.4 AUTOMATED AIRCRAFT AND HUMAN PERFORMANCE

When increased automation of aircraft was first envisaged, Wiener (1993) argued that effectiveness, efficiency and flight safety would benefit substantively. This implied that there were financial implications for airline companies that chose not to operate modern equipment. It made business sense for both entrepreneurs and governments to invest in technology (Bainbridge, 1983). In consequence of these changes, it was predicted that over time, a number of pilots would have no choice but to transition from analogue to digital flight control systems (Bent, 1996; Chambers & Nagel, 1985).

The current body of literature points out that new human factor issues that are unique to technological changes are increasingly being raised by airline managers, accident investigators, civil aviation regulators and human behaviour experts (CAA, 2011; NTSB, 2009; Parasuraman & Riley, 1997). Poor interface design, human complacency, over-reliance on automation, a loss of manual flying skills, and pilots' general lack of understanding of design intentions and system logic are some of these human-related concerns.

One regulator's report into a serious incident involving an Airbus A340-300 aircraft in Johannesburg cited pilot training and a lack of understanding of the system design as direct and significant contributory causes of the incident (CAA, 2011). On further inspection of the training material (Airbus, 2011b) associated with this particular accident, it was determined that the recommendation to use a FBW tool linked to the side-stick control during lift-off could be confusing to pilots. A Maltese cross on the primary flight display was in no way correlated with the actual elevator or aileron flight control surface position, and was thus only an indication of side-stick deflection, which led to pilots' confusion on the flight deck. This illustrates how a relatively trivial component of the system (not anticipated as a problematic area to the design engineers) can lead to a critical breakdown in aircraft understanding by the pilots and subsequently to a serious incident.

The manufacturer has since updated the software logic to remove a part of the indicator that caused the human factor problem in this particular incident.

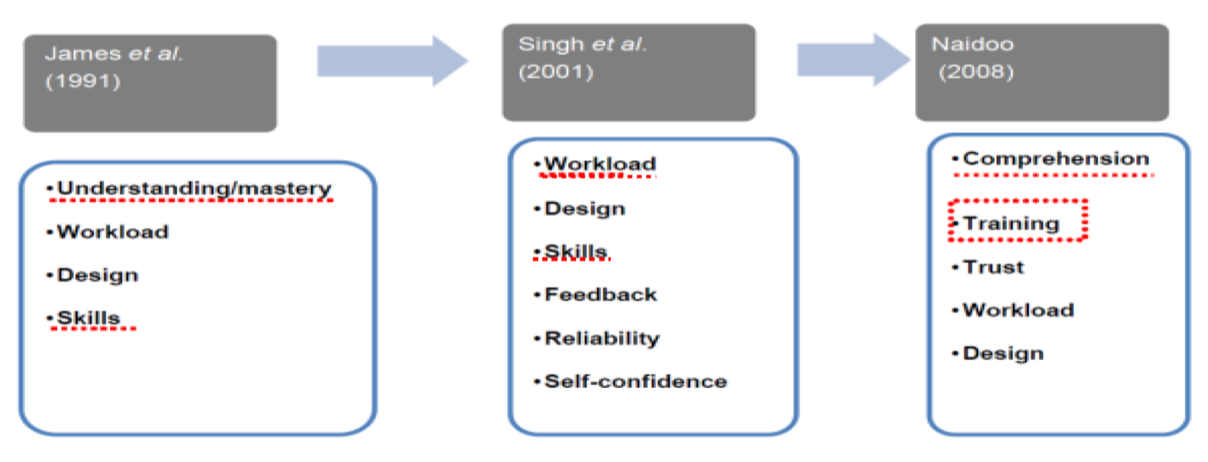
Furthermore, this incident draws attention to the fact that no amount of ground testing can ever cater for all possible permutations of the human-machine dynamic in actual flight. There will always be some combination of human-machine circumstances that may result in unforeseen outcomes (Rigner & Dekker, 2000).

In another serious incident involving complex peripheral subsystems, misidentification from a faulty computer processed fuel pump resulted in engine thrust loss (Hradecky, 2011). The monitoring computer of this aircraft received anomalous information and passed this information on to the crew. Ambiguity in the human-machine interface led the crew to shut down fuel pumps that were, in fact, functioning correctly. Fortunately, the crew were able to pick up the error after an engine rolled back, and they then disregarded the recommendations from the monitoring computer system, and were able to restore engine thrust. In this case, the human operators had sufficient knowledge and understanding of the system to avert disaster. These, and other similar incidents, serve to highlight the impact of automation on human performance.

Empirical research into airline pilots' experiences with advanced automated aircraft (the human-machine interface) spanning the last two decades shows that an undesirable human-machine relationship may be gradually emerging (Mitchell *et al.*, 2009). A comparison of three important automation-related aviation surveys confirms empirically that automation *training* is the new factor extrapolated from the item correlates. The emergence of this factor opens up a new research path in aviation psychology.

An analysis of the advanced automated aircraft training climate will provide some of the much-needed new knowledge in this direction. Figure 8 depicts the trend of factors revealed by the three surveys conducted over the last 20 years. The factors highlighted in the figure suggest the emergence of a mismatch between systems knowledge and design intentions – the training-related areas are clearly marked.

Figure 8: Trends in three aircraft automation surveys



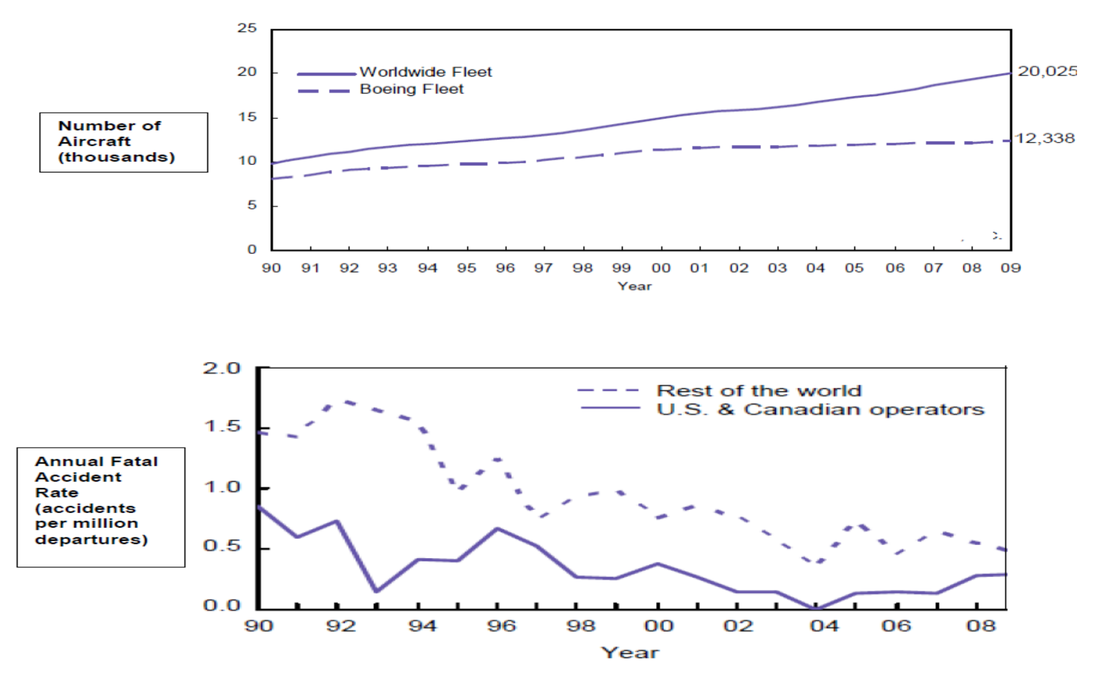
Source: Adapted from Mitchell *et al.* (2009)

The three surveys were used to collect empirical evidence about emerging issues related to advanced automated aircraft. Clearly, Figure 8 shows that the factor of *training* appears to have become a more important issue over the years as aircraft become more complex.

There are a number of general concerns about human beings' interaction with technology, as well as specific concerns, such as systems design and implementation. Examining the approximately 92 distinct automation issues affecting human operators substantiates this statement (Lyall & Funk, 1998; Parasuraman & Byrne, 2002; Parasuraman & Riley, 1997). Detailed discussions of these issues fall beyond the scope of this literature review – suffice it to say that the web database http://flightdeck.ie.orst.edu/ contains a quality discussion on specific topics which will be valuable in any research dedicated to advanced aircraft issues (Lyall & Funk, 1998:288). The database was formulated on the basis of evidence from accident and incident reports, experiments, surveys and other studies.

For the purposes of the present study, a review of the database was deemed essential for understanding the human-machine dynamic at a specific human-factors level. Reviewing the database also served as a guide for the development of various questionnaire items for the initial phase of the instrument development. The issues are verified and validated with empirical evidence from both qualitative and quantitative reports (see Tables 4 and 5). Independent research by the big aircraft manufacturers Airbus (2011b) and Boeing (2009) found that nearly 20% of human error accidents are directly related to the human-machine interface, and specifically to human interaction with on-board automation. Figure 9 compares the overall accident rate with the manufacturing rate of modern western-built commercial jet aircraft (over thirty tonnes).

Figure 9: Aircraft production versus accident rate



Source: Boeing (2009)

The study conducted by Boeing (2009) found that, although the use of advanced western-built commercial jet aircraft weighing over thirty tonnes has increased within the last two decades, the moving average accident rate for these aircraft has decreased substantially (see Figure 9). This decrease is ascribed mainly to advances in design and technology. However, even with exceptional technology, it appears very difficult to achieve the elusive zero accident rate, in part due to the adverse effects and contributory impact from the human factor (Dutch Safety Board, 2009; Machin & Fogarty, 2003). Although the accident rate is far lower with modern automated aircraft than with more traditional analogue-based aircraft in general terms; psychologists, researchers, pilots and aviation safety experts are still concerned about the breakdown in knowledge loops which lead to aircraft incidents and accidents in modern digital aircraft (Johnston *et al*., 1995; Machin & Fogarty, 2003).

Lyall and Funk (1998:291) list the following concerns regarding the human machine-interface, after quantification from various sources:

- pilots may place too much trust in automation;
- there has been a loss of manual flying skills; and
- pilot interface systems may be too complex for the average pilot.

Paralleling the aforementioned concerns, in an effort to pinpoint the issues, regulators and scholars in the field have also attempted to summarise the issues that affect flight crews' optimum use of automation (Airbus, 2011b; Helmreich, 1987; Parasuraman & Riley, 1997; Risukhin, 2001). The following issues were used to guide the present study toward an encompassing measurement construct:

- *being intimidated,* which may prevent pilots from taking over from the autopilot until a very late stage;
- *overconfidence in, or overreliance on, the autopilot,* which may make pilots delegate too many tasks to the computer too often;
- *inadvertent arming or engagement* of an incorrect mode;
- *failure to cross-check* and verify the armed automation mode;
- selection of *incorrect automation targets*;
- *a lack of discipline* in confirming selected automation targets on primary reference instruments;
- *a preoccupation with flight management programming* during critical phases of flight, with a consequent loss of situational awareness;
- *a lack of understanding* of the automation mode transition process and mode reversions, resulting in mode confusion (misunderstanding the autopilot); and
- *poor crew resource management* (CRM) practices, resulting in inadequate task sharing and monitoring.

2.4.1 The impact of human factors on aircraft safety

“Present technology is characterized by complexity, rapid change and growing size of technical systems. This has caused increasing concern with the human involvement in system safety” (Rasmussen, 1990:449). Some research suggests that organisational human error, and specifically pilot error has contributed to nearly half of all accidents involving western-built commercial aircraft above 30 tonnes (as depicted in Table 4). Furthermore, Table 4 illustrates comparatively the impact of human behaviour as a cause of accidents with various other non-human causes. Adverse weather-related phenomena and mechanical faults are regarded as threats emerging within the operating environment and which subsequently manifest into on-board crew related errors (Helmreich, 2002). It has always been the intention of advanced aircraft manufacturers to improve automation systems so as to assist pilots in dealing with non-human sources of environmental threats. Paradoxically however, some studies have found increased technology related human error such as automation complacency, when dealing with external threats such as adverse weather or substandard navigational facilities (Parasuraman & Byrne, 2002).

Table 4: Accident statistics for western-built commercial aircraft above 30 tonnes

Cause	1950’s	1960’s	1970’s	1980’s	1990’s	2000’s	All
Pilot Error	41%	34%	24%	26%	27%	30%	29%
Pilot error caused by weather	10%	17%	14%	18%	19%	19%	16%
Pilot Error caused by mechanical issues	6%	5%	5%	2%	5%	5%	5%
Pilot error total	**57%**	**56%**	**43%**	**46%**	**51%**	**54%**	**50%**
Other human related errors (air traffic controller errors, improper loading of aircraft, fuel contamination, improper maintenance procedures)	2%	9%	9%	6%	9%	5%	7%
Adverse weather	16%	9%	14%	14%	10%	8%	12%
Mechanical failure	21%	19%	20%	20%	18%	24%	22%
Sabotage	5%	5%	13%	13%	11%	9%	9%
Other	0%	2%	1%	1%	1%	0%	1%

Source: Adapted from National Transportation Safety Board (2009)

According to Helmreich (2002:18), human error can generally lead to an aircraft incident or accident if not “trapped” early. Therefore, any critique of human error requires some reference to aircraft incidents or accidents. For the discussion in this section, it becomes necessary to clarify the difference between an aircraft accident and an aircraft incident. The International Civil Aviation Organisation (2001) defines an *accident* as any occurrence associated with the operation of an aircraft where a person is fatally or seriously injured (except from natural causes, self-inflicted or from others, or injuries to stowaways), the aircraft sustains structural damage affecting flight performance, or requires major repair, or the aircraft is missing or completely inaccessible. An *incident* on the other hand, is regarded as an occurrence with the potential for affecting safety, which can lead to an accident.

Table 4 clearly shows how human factor issues, specifically those involving pilot error, should be a concerning factor for accident/incident investigators, regulators and operators. Pilot error can be regarded as a mistake, omission, commission, lapse, negligence or faulty judgment on the part of the pilot, which may lead to an incident or accident (Risukhin, 2001). Parasuraman and Byrne (2002) pointed out that the introduction of highly advanced commercial aircraft during the 1990’s have contributed to the increased human factor issues related to the accident rate, and this can be clearly seen from the descriptive statistics illustrated in Table 4.

A typical aircraft accident investigation, specifically those involving elements of the human factor, may last up to five to ten years and can still remain inconclusive even after many years of painstaking investigation, due to the complexities and controversies involved (National Transportation Safety Board, 2009). Investigating human error in aircraft accidents is a contentious issue, specifically when it involves highly advanced technology, as many players are subsequently associated; such as the operators, designers and managers, therefore not just the pilots (Rasmussen, 1990; Rigner & Dekker, 2000). The complexity arises from the complication in the relationships and interests between the various stakeholders involved in such an accident. For instance, manufacturers may want to blame the pilots; the airline companies may want to blame the aircraft; passengers and the families of the flight crew may want to blame managerial decision-making. Moreover, human factor-related incidents and accidents contain a fair amount of subjectivity when gross negligence or intentional non-compliance with procedure is not the case. Pinpointing the exact nature of human error on the flight deck when it involves interaction with high automation is difficult, particularly because automation was intended to relieve pilot workload (Singh *et al.*, 2005).

To put it simply, air crash investigations highlighting human factor issues relating to technologically advanced aircraft are complex because these investigations involve psychology and examining the two main advanced aircraft components. In analysing such accidents, incidents or mishaps, it was concluded that in general a distinction can be made between the automated systems affecting the pilot in the flight deck, such as electronic flight instrument systems (EFIS), the flight director (FD), navigation, or the flight control unit (FCU), on the one hand, and the advanced airframe mechanical subsystems, such as flight controls (FBW), hydraulics, electrics or electronic engine control, on the other hand. Table 5 and 6 was drawn up to highlight such distinction. These tables also summarise incidents involving some older analogue type aircraft automation systems, together with the more relevant advanced automated aircraft systems for more enlightening contrast. Although improvements were made in some systems after aircraft accidents, similar human factor and training problems continue to haunt modern digitised aircraft.

Table 5 shows that many of the incidents and accidents concerning automated flight deck systems involve the pilot's misinterpretation of the flight computer state. The phenomenon is called mode confusion. According to Parasuraman and Byrne (2002), the problem arises when there is a mismatch between a pilot's mental model of reality and the actual aircraft situation as interpreted by the flight computers. For instance, Table 5 goes on to list many accidents resulting from vertical mode confusion. In these cases, the pilots assumed that the aircraft would maintain a particular descent trajectory; however, in reality, many pilots continually select an incorrect autopilot mode, one of the major problems in the advanced flight deck. The aircraft however, would only perform as demanded by the pilot, resulting in a dangerous set of circumstances, such as idle thrust very close to the ground. Aircraft manufacturers therefore stress the importance of continually maintaining a high level of situational awareness through correct and accurate interpretation of the flight computers and subsequent aircraft state. For example, a core Airbus golden rule states that pilots should know (understand and interpret) the Flight Mode Annunciator (FMA) at all times (Airbus, 2011b). The FMA is possibly one of the most important indications of the current state of the aircraft in a glass flight deck and should be considered a primary instrument (Funk & Lyall, 2000).

Table 5: A chronological list of automation incidents and accidents related to the flight deck

Automated aircraft flight deck systems					
Year	**Location**	**Aircraft type**	**Operator**	**Description of incident or accident**	**System(s) involved**
1972	Miami	L-1011	Eastern Airlines	Loss of situational awareness after an inadvertent autopilot disconnection.	ALTITUDE HOLD
1973	Boston	DC-9-31	Delta Airlines	Pilots' preoccupation with questionable flight director led to a loss of situational awareness.	FLIGHT DIRECTOR
1988	Gatwick	A320	Air France	Vertical mode confusion.	FLIGHT CONTROL UNIT
1989	Boston	B767	Unknown	Vertical mode confusion.	FLIGHT CONTROL UNIT and FLIGHT DIRECTOR
1990	Bangalore	A320	Indian Airlines	Vertical mode confusion.	FLIGHT CONTROL UNIT
1991	Moscow	A310	Interflug	Inadvertent autopilot disconnection leading to confusion and loss of control.	ELECTRONIC FLIGHT INSTRUMENT SYSTEM
1992	Strasbourg	A320	Interair	Vertical mode confusion.	FLIGHT CONTROL UNIT
1993	Tahiti	B744	Air France	Inadvertent autopilot disconnection and vertical mode confusion.	NAVIGATION MODE
1994	Toulouse	A330	Air France	Unexpected altitude capturing during a simulated engine failure.	NAVIGATION MODE
1995	Connecticut	MD80	American Airlines	Inadvertently descended below minimum altitude.	NAVIGATION MODE
1995	Cali	B757	American Airlines	Incorrect input into the flight management computer resulting in aircraft impacting terrain.	NAVIGATION MODE
1996	Puerto Plata	B757	Birgen Air	Loss of control.	ELECTRONIC FLIGHT INFORMATION SYSTEM

Source: Adapted from National Transportation Safety Board (2009); Helmreich, 1987; Parasuraman & Riley, 1997; Risukhin, 2001

Table 6: A chronological list of automation incidents and accidents related to airframe subsystems

Automated aircraft mechanical subsystems					
Year	**Location**	**Aircraft type**	**Operator**	**Description of incident or accident**	**System(s) involved**
1984	New York	DC10	Scandinavian Airlines	Overran runway.	POWER PLANT
1985	San Francisco	B747	China Airlines	Inappropriate control of engine failure using the autopilot system.	POWER PLANT and ELECTRONIC ENGINE CONTROL
1988	Habsheim, France	A320	Air France	Loss of situational awareness in flight envelope.	FLY-BY-WIRE CONTROL SYSTEM
1989	Helsinki	A300	Kar Air	Inadvertent activation of Go-Around mode.	ELECTRONIC ENGINE CONTROL
1999	Warsaw	A320	Lufthansa	Overran runway.	POWER PLANT mode logic
1994	Hong Kong	A320	Dragon Air	Incorrect flap setting.	FLAPS MANAGEMENT SYSTEM
1994	Nagoya	A300	China Airlines	Aircraft inadvertently stalled on final approach.	ELECTRONIC ENGINE CONTROL
1994	Manchester	B757	Britannia	Inadvertent stall situation, recovered.	POWER PLANT and ELECTRONIC ENGINE CONTROL
1994	Paris	A310	Tarom	Aircraft inadvertently stalled then recovered.	POWER PLANT and ELECTRONIC ENGINE CONTROL
1994	Indiana	ATR72	American Eagle	Lack of knowledge in flight surface de-icing system led to inadvertent stall.	DE-ICING SYSTEM
1995	Bucharest	A310	Tarom	Aircraft entered a spiral dive situation.	ELECTRONIC ENGINE CONTROL
2008	Sao Paulo	A320	Tam	Overran runway after confusion with auto thrust.	ELECTRONIC ENGINE CONTROL
2009	Schiphol, Netherlands.	B738	Turkish Airlines	Inadvertent aircraft stall on final approach after thrust auto reduced to flight idle.	ELECTRONIC ENGINE CONTROL and AUTO THRUST
2009	Atlantic ocean	A330	Air France	Aircraft stalled after loss of flight information and autopilot.	FLIGHT CONTROL COMPUTER

Adapted from National Transportation Safety Board (2009); Helmreich, 1987; Parasuraman & Riley, 1997; Risukhin, 2001

Table 6 depicts a number of power plant- or engine-related problems. The full authority digital engine control (FADEC) system of many advanced aircraft is both a complex and highly efficient system, one which has allowed modern aircraft to generate profits for the companies that own these aircraft. It allows aircraft to be flown with such high precision, that fuel savings have increased substantially over the last 10 years. However, the high complexity of the system has also resulted in new and previously unheard of human factor errors when the system breaks down (Cockburn, 2007; Rouse & Morris, 1987).

2.5 AIRLINE PILOT TRAINING

The changing role of human beings' relationships with technology is partly a result of the blurring of boundaries between the technical and non-technical expertise required to perform effectively (Funk & Lyall, 2000). Rigner and Dekker (2000:318) suggest that a modern aircraft pilot (as opposed to a traditional "stick-and-rudder" pilot) is a proactive manager of a complex system. A modern airline pilot is required to resolve complex automation problems involving supervising, programming, monitoring, and cognitively deciding on tactics or strategies incorporated in an array of complex computers (Mosier *et al.*, 2007). The paradigm shift requires airlines to rethink their training regimes.

Caro (1998) contends emphatically that a basic requirement to meet the need for precision in flight training methods and syllabi is a systematic analysis of piloting tasks or the required competencies in a changing learning environment. Furthermore, it is argued that "imprecisely defined aircrew training programs cannot demonstrate the relevance and adequacy of their course content with respect to known training requirements and, therefore, [organisations] might be judged culpable in the event of errors committed by aircrews they trained" (Caro, 1988:249). Airlines have absorbed many of the aspects of training pilots on complex, advanced aircraft, such as transition training. Scholars in the subject are beginning to find it increasingly difficult to gain access to potential samples of airline pilots undergoing training on advanced aircraft, because airlines tend to limit access to outside researchers (Funk & Lyall, 2000; Rigner & Dekker, 2000).

Large airline organisations are commonly referred to as legacy carriers. Such carriers are inclined to employ only qualified and experienced pilots who are recruited to complete an organisationally structured training course to operate commercial aircraft in accordance with precise measurement outcomes (Taylor & Emanuel, 2000). Instructors tasked with training such airline pilots emphasise the technique of integrating known learning (in other words, how to fly older aircraft) with the unknown (how to fly modern digitised aircraft) by drawing on pilots' experiences (Bent, 1996; South African Airways, 2007). This strategy has proved fruitful in mitigating the complexities of modern aircraft training. In support of the strategy, the minimum experience levels required to join a legacy carrier are therefore extremely high. The effects of this recruitment policy are clearly noticeable from the demographics in the present study's sample frame.

2.5.1 Airline training strategies

Two types of initial training for airline pilots are commonly differentiated in the literature: pilots have either *civilian* or *military* training backgrounds (Taylor & Emanuel, 2000). Although military-trained pilots receive less team-based training during their initial flight training, they are considered highly experienced and skilful, and are thus much sought after by commercial airline companies (Bent, 1996). It is common knowledge in the aviation industry that military candidates complete an easily verifiable precision-based and highly structured training course (Andrews & Thurman, 2000). Powerfully regimented training programmes are very difficult to replicate in non-military settings (Caro, 1988).

In training an advanced aircraft pilot, it is the airline organisation that is generally responsible for a candidate's transition onto a new aircraft. This transition consists of three broad components. This multidimensional approach to training stratification consists of a theoretical learning part, a flight simulator training part and a route (or actual flying) part (Moore, Lehrer & Telfer, 1997). Flight simulation is possibly the most critical pedagogical aspect of an airline pilot's training (Pasztor, 2009). Since using a flight simulator is a critical and legal requirement for training pilots in unusual and emergency type scenarios or situations, it is discussed separately, in Section 2.5.3.

Converting trainees and maintaining their competence on a new aircraft type is generally accepted as the responsibility of the airlines that employ pilots. This conclusion is substantiated from a review of the literature which shows how empirical research consistently suggests that organisational level gaps in knowledge, misunderstandings and misconceptions, are responsible for a fair proportion of the problems and failures associated with airline pilot transition training (Lowy, 2009; Lyall & Funk, 1998; Parasuraman & Riley, 1997). Recently, the NTSB's severe incident report into an advanced aircraft issue found that the company-designed training material contained items that conflicted with some of the practices recommended by the manufacturer (Hradecky, 2011). Therefore, it is clear that misunderstandings and misconceptions of complex systems can frequently originate at the organisational or macro level. This invariably leads to significantly inappropriate behaviour at an individual (pilot) level at an operational level (Patrick, 2002). Scientific examination and research into the learning environment may give investigators insight into these and other problems.

Maintaining a competitive advantage in an industry with very narrow margins implies that businesses must, and do, invest in new technology (Australian Transport Safety Bureau, 2007). Funk and Lyall (2000) found that the technological complexities of automated aircraft mean that organisations will continuously teach new skills to pilots or (re)train pilots and that "increased levels of automation with the advent of the *glass cockpit* have resulted in substantial changes in the way civilian aircrews are trained" (Taylor & Emanuel, 2000:18, own emphasis).

Training can be an expensive exercise for an airline business, and therefore it should be regarded as an investment for the long term. The typical airline organisation would therefore invest in various critical resources to meet current and future training demands. These resources, according to Caro (1988), include assets such as personnel to conduct the instruction, operational aircraft, simulators (or other aircraft representations), printed or graphic media, classrooms or practice areas, and variations of other specialised aids or devices. The resources chosen all depend on the tasks that pilots need to master. Some researchers argue that a paradigm shift is needed in how organisations train pilots for advanced automated aircraft in order to meet the needs of humans in understanding complex computerised systems (Funk &

Lyall, 2000; Naidoo, 2008; Rigner & Dekker, 2000). Therefore, it may be necessary to change, enhance or develop new training materials for the advanced aircraft.

Previous research has found that airlines are slow to make relevant changes to training methodology. For example, combating “automation bias and complacency”, particularly with very experienced advanced aircraft pilots, is an area of training, which airline organisations often seem to neglect (Mosier *et al.*, 2007:301; Naidoo, 2008:110). The emphasis is still placed on training aspects related to an earlier era of aircraft, such as engine failure and manual handling skills, which include trainees' demonstration of conducting accurate steep turns (banking the aircraft outside of its normal operational envelope). Some authors argue that the reliability of aircraft engines is such, that less than 1% of airline pilots will experience an actual failure on the line, so more emphasis should be placed on handling automation failure rather than on manual control of an engine failure (O'Hare *et al.*, 1994). Some authors suggest that it is bizarre that a pilot of a highly advanced aircraft should have to demonstrate to the regulators an ability to fly the aircraft outside its normal operating range in order to receive a licence (Lyall & Funk, 1998; Poprawa, 2011). For example, pilots may still have to prove that they can accurately perform a steep turn, which pushes the aircraft beyond its normal bank angle (when such a manoeuvre is accomplished fairly easily in modern fly-by-wire aircraft such as the Airbus family, and therefore demonstrates very little skill). These and other similar debates illustrate the level of uncertainty in training pilots to fly advanced aircraft.

2.5.2 Models of airline instruction

In response to the systemic nature of training and instruction in a modern airline, Spector and Muraida (1997) suggest that models of Instructional Systems Development (ISD) or Systems Approach Training (SAT) be consulted in planning multidimensional learning environments. The systems methodology of pilot training was reviewed based on the systemic approach to the research construct (discussed in Chapter 3). The systems training method is thought to maintain the precision required for advanced aircraft pilot training (Bent, 1996). These models are used extensively and have served the advanced aircraft training community well over the last decade (Panda, 2003). Operationalization of the construct to measure the advanced aircraft training climate in the current study was similarly multidimensional. Andrews and Thurman (2000) argue that aviation training organisations may justify deviating from the prescriptions of a particular multidimensional structured ISD model only for logical reasons. The current literature suggests that the effects of such training methodological decisions are not yet fully understood, but anecdotal evidence suggests that following an ISD model too closely can in itself present some dangers – for example, Hradecky (2011) reports that crew can be incorrectly trained (albeit to company standards) to deal with complex aircraft problems.

Table 7 is presented as a synthesis of some important models suggested by various experts whose work contributed to building the initial research framework.

Table 7: Chronological synthesis of Instruction Systems Design models (ISDs)

Source	Model Description
Spector and Muraida (1997:67)	1. Conduct a needs analysis. 2. Design the course. 3. Produce the programme. 4. Implement the course. 5. Continually maintain the course.
Pohlman and Fletcher (1999:297)	1. Analyse the job by asking what knowledge skills, outcomes, and attitudes are to be produced. 2. Design the instruction and devise the instructional interaction. 3. Produce, develop and prepare the instructional materials. 4. Install and implement an appropriate training system. 5. Evaluate, verify and validate the instruction.
Patrick (2002:439)	1. Identify the training needs and required tasks that need to be trained from qualitative and quantitative accident/incident analysis. 2. Design the training based on appropriate psychological principles and theories to promote motivation and ensure learning transfer. 3. Evaluate whether the training programme has actually achieved its intended objectives.
Panda (2003:129)	1. Analyse the training requirements. 2. Design the training programme. 3. Develop the course. 4. Implement the training programme. 5. Control the training programme.

An examination of the models in Table 7 reveals that the frameworks suggested all contain some version of the steps commonly employed in systems engineering designs (Panda, 2003). The ISDs are thus a scientific approach to training. The sample used in the current study consisted primarily of organisations that conform to the basic ISD framework. The use of an ISD model in an aircrew-training environment is based on psychological, philosophical and pedagogical orientations (Pohlman & Fletcher, 1999). Therefore it makes intuitive sense that an ISD will generate competent pilots more often than not. For these reasons, analyses of such frameworks were deemed pertinent in the present research approach in designing an appropriate measurement construct.

Approaching an analysis of the aviation environment in a scientific and systemic manner can ensure that operators are as objective as possible in training, assessing skills, and overall in qualifying competent pilots.

2.5.3 Flight simulator training

Meister (1999) contends that the pilot and the aircraft are fundamentally and critically interrelated as a system. Therefore, indirect aviation measures of human performance can be extracted from the state of the aircraft, or in a training situation, of the experiences in a flight simulator (Bonner & Wilson, 2002). For instance, Dahlstrom and Nahlinder (2006) suggest that there is immense value in using flight simulation as a source of information to improve basic civil aviation training. Therefore, a core modern training device employed by airline organisations for the structured training of advanced aircraft pilots is the flight simulator, or synthetic training device (FSTD). Furthermore, FSTDs are a legal requirement for training pilots engaged in any commercial flight operations at airline organisations (Civil Aviation Authority, 2011).

The first aircraft flight simulator was built before World War I to simulate the Antoinette monoplane (Rolfe & Staples, 1986). Flight simulators and various synthetic training devices will continue to play a critical role in training the modern or advanced aircraft airline pilot, today and for the recognisable future (Magnusson, 2002). This component has in turn made a significant impact on how pilots experience their overall aircraft training (Telfer, Moore & Farquharson, 1996). The core intent in the usage of the aircraft flight simulator is possibly to ensure that the simulation should accurately mimic reality as closely as possible (Howell & Fleishman, 1982). However, Dahlstrom and Nahlinder (2006) found sufficient evidence that the mental workload to perform in a flight simulator is far less than that required in actual aircraft flight, which suggests that there may be a mismatch between simulation and reality. Flight simulator realism nonetheless, is based on a basic simulation structure in three parts, as proposed by Rolfe and Staples (1986:4):

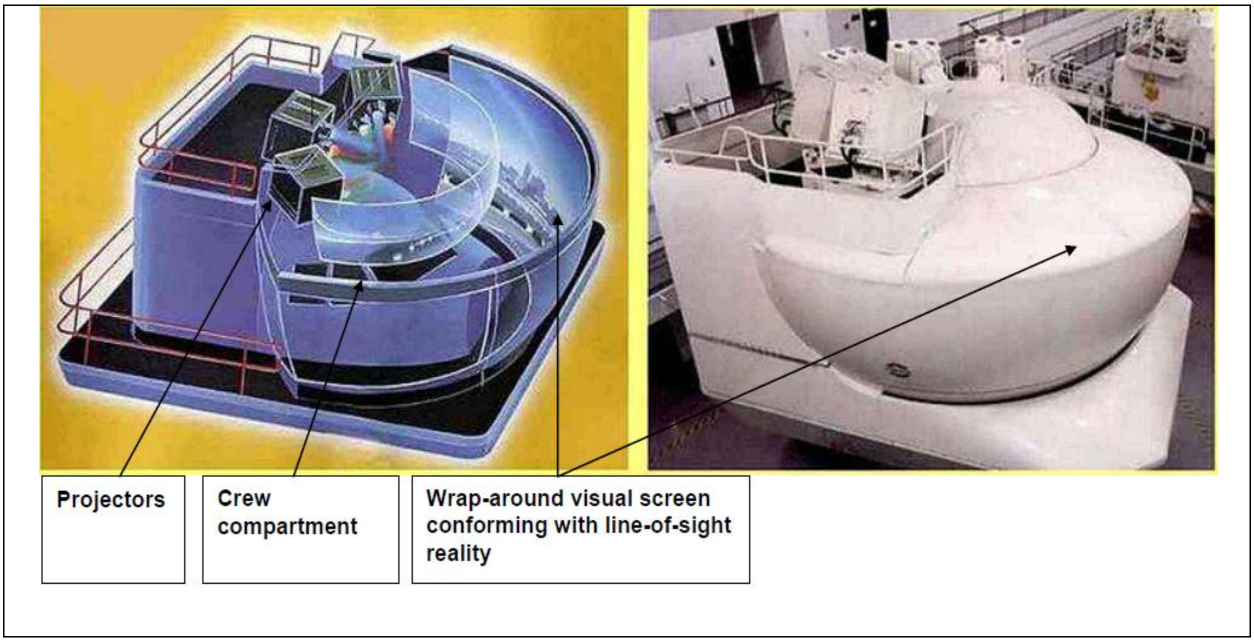
- a model of the system to be simulated;
- a device through which the model is implemented; and
- an applications regime to satisfy the combination of the first two elements in such a way as to meet the training objectives.

If any of the aforementioned components are missing, it may be expected that the gap between simulation and reality should increase. More research may be needed to compare flight simulation to actual aircraft flight in order to determine precisely the level of disconnect. This may have consequences for programmes such as the MPL (Multi Pilot Licence). The MPL entails reducing a trainee's actual aircraft flight experience by replacing it with flight simulation time as a method to expedite pilot training (ALPA-SA, 2011).

Pilots are permitted to practise emergency and unusual situations only in an aircraft flight simulator, and not in the actual aircraft, due to safety considerations (Civil Aviation Authority, 2011). The details of these procedures for licensed pilots and training organisations are found in the primer stipulating the international civil aviation regulations and technical standards for all signatories of the Chicago Convention (International Civil Aviation Organisation, 2011), which includes the majority of countries operating modern western-built commercial jet aircraft.

The theory of common elements and transfer surface forms the basis of the simulation concept. Projectors and wrap-around visual screens in modern flight simulators attempt to mimic reality closely (Rolfe & Staples, 1986). Technologically modern FSTDs employ high resolution, computer-generated colour images, operating in multiple degrees of freedom (Howell & Fleishman, 1982). The realistic appearance of the synthetic flight deck (Figure 10) suggests that operational transfer will occur to the extent that there should be definite commonalities between the simulator and the aircraft (Thorndike, 2007). Evidence from some experiments confirmed a strong correlation between elements, or features, contained in the simulator and the features of the actual equipment, suggesting that a transfer of training is positive or high from an advanced FSTD (Bonner & Wilson, 2002; Caro, 1988). Therefore, airline operators and instructors are extremely confident about the level of training received by their pilots and its transferability to the actual aircraft. There may be a strong financial incentive for organisations to use modern simulators, because operating the actual aircraft to obtain licence ratings is no longer a prerequisite. Indeed, confidence in the accuracy of the FSTD is so high that regulators allow a pilot to obtain a licence for an advanced aircraft without having flown the actual machine. The level of psychological connection or disconnection experienced by the trainee pilot between an FSTD and the actual aircraft is therefore an important area for further investigation.

Figure 10: Modern flight simulator training device



Source: Adapted from Stevens and Lewis (2003) and Strachan (2011)

The crude combination of appropriate levers and linkages in the first monoplane flight-training simulator has since evolved significantly, into a highly sophisticated synthetic device in an attempt to replicate reality as accurately as possible (Figure 10). Advances in simulation techniques provided by computer, materials and engineering technology attempt to recreate a supposedly seamless integration between the FSTD and actual aircraft fight (Magnusson, 2002). Such training is commonly referred to as zero flight time training (Stevens & Lewis, 2003). However, contradictory findings suggest that simulator transition training is not and can never be a completely seamless exercise, as there are always gaps between reality and simulation (Singh, Sharma & Singh, 2005).

One of the main problem areas, which result in such mismatches, stems from the fact that trainees can pre-empt an emergency exercise in the flight simulator, whereas this is not the case in actual aircraft flight. Therefore, it was found that the heart rate and mental workload of pilots in simulated exercises do not completely correlate with that in actual aircraft flight (Dahlstrom & Nahlinder, 2006). Nonetheless, Go, Bürki-Cohen and Soja (2003) found that the full motion simulator did indeed make a statistical difference to the evaluation of pilots; however, it played only a middling role in actual pilot training. Perceptions of training in a flight simulator may then be affected. In other words, because all aspects of flight training for an advanced aircraft cannot take place in the actual aircraft, it may be concluded that simulators may influence pilots' training experiences either negatively or positively. Therefore, one of the secondary goals in the current research was to gain further understanding of the phenomena associated with the use of flight simulator training devices.

2.5.4 Pilot route training

Caro (1988) and later, Go, *et al.* (2003), found a significant relationship between training utilising an FSTD, to the training gained from actual aircraft flight. In order for an airline company employing an advanced aircraft pilot to gain sufficient evidence that this transition (from simulation to reality) has indeed been successful, route training is mandatory for a new pilot on aircraft type (South African Airways, 2007).

Route training generally consists of flying a predetermined number of sectors in normal operations to ascertain the level of competence the trainee has achieved from the FSTD (SAA, 2007). The candidate is required to complete a number of tasks in the real aircraft with an instructor present. Exercises such as landing the aircraft with varying flap settings, cross-wind approaches, landings and take-offs in adverse weather and other normal operations are completed by the trainee pilot, where after the pilot is deemed fit to operate the aircraft in normal line operations.

Airline training organisations are well aware of the divide that may exist between flight simulation and actual aircraft flight (Thorndike, 2007). This has made actual aircraft flight training in the form of route training, a mandatory requirement in qualifying an advanced aircraft pilot. With this in mind, it was therefore necessary to probe trainees' perceptions in terms of both their experiences in the synthetic training device and in actual aircraft flight.

2.6 CONCLUSION

The literature review has highlighted the current lack of empirical academic knowledge of the psychology associated with advanced aircraft training. More objective scientific research is needed to answer the following question posed by Barker (2011:4): "Is automation error going to be the new human factors contribution to accident statistics?" Many analyses associated with advanced aircraft training were inconclusive regarding the psychological attributes that affected such training. The current study proposes the proposition that technological complexities in advanced automated aircraft have resulted in a shift in the role being played by human beings within the human-machine dyad. The competence required from pilots in respect of both their technical and non-technical abilities when operating advanced automated aircraft suggests that a paradigm shift is needed in the way organisations view their training regimes.

The chapter has mentioned the economic and safety motivations behind increased aircraft automation and some of their implications. Any advances in aircraft-related technology need to go hand in hand with new or additional training requirements and resources such as synthetic flight simulation devices. Effective and efficient training is a critical component in enhancing overall flight safety by mitigating the effects of human factor issues. It was found that the literature supports the premise that changes to the flight deck have resulted in unforeseen human factor issues. Because these issues are not easily designed out of aircraft, a close relationship is required between research psychologists and engineers.

It appears that airline organisations are slow to adapt to changes in the external environment, specifically in terms of training paradigms. Traditional transition training has made it difficult for scholars to assess fully the multivariate phenomena present in the systemic and rapidly evolving aviation environment. More importantly, the literature review reveals a real need for an appropriate psychological assessment scale to measure the training aspects of constructs associated with training pilots to fly advanced automated aircraft.

Determining what constitutes a suitable training climate for technologically complex systems may make it possible to understand the psychological and behavioural components of exactly what new knowledge acquisition is, and the subsequent transfer of learning into safely managing advanced machinery. Hence, the chapter examined, in behavioural terms, advanced flight deck automation. This was followed by a discussion of how airline pilots possibly undergo training for new technology. The review suggests that the synthetic flight training device or FSTD (flight simulator) is critical in teaching the modern airline pilot how to perform many tasks. This chapter has examined the complexities associated with an advanced aircraft training environment. The literature shows that the introduction of human beings into such an intricate environment induces a new dynamic, created by both psychological and behavioural components. In order to scientifically measure phenomena related to these components, it is necessary to operationalize an appropriate construct (see Chapter 3). The scientific measurement of constructs provides an avenue for specific and focused findings, rather than general or global discussion. Therefore, the next chapter presents and discusses the organisational, instructional and individual aspects required for developing appropriate theoretical models to meet the study’s objectives.

CHAPTER THREE:

LITERATURE STUDY –
THE ADVANCED AIRCRAFT TRAINING CLIMATE

3.1 INTRODUCTION

> *Education consists mainly of what we have unlearned.*
> (Mark Twain, 1898, cited in Keane, 2005:204)

> *Any sufficiently advanced technology is indistinguishable from magic.*
> (Clarke, 1984:27)

The realm of advanced computing can cause confusion and often uneasiness amongst the general population (Moore, 2003). Linking two very complex areas – advanced technology and training – can therefore be a daunting research task. The purpose of this chapter is to introduce the organisational training climate in the context of advanced automated civilian commercial aircraft. A multidisciplinary approach was therefore adopted to interpret the vast array of literature related to the topic. Evidence was sourced from the field of organisational behaviour (which focuses on multiple levels of analysis) and from contemporary theories originating in clinical and industrial psychology, as well as from literature on aviation human factors, advanced computerisation, aircraft accident investigation principles, and learning and training (educational psychology and sociology). This chapter encapsulates work from the existing body of knowledge to underpin and contextualise the development of a tool to measure perceptions of the advanced aircraft training climate.

The impact of technology on both the latent and overt behaviour of airline pilots has been thoroughly documented, researched and mapped, but only up to a point (Parasuraman & Riley, 1997; Research Integrations Inc., 2007). Thus far, there has been little scientific analysis of the *environment* in which pilots *acquire* and *adapt* new knowledge related to working with advanced automated commercial aircraft.

Science starts with good definitions (Carston, 2002). Difficulties in science arise when what an author intends to say diverges from what is actually understood in a particular context. Problems in research reporting arise where there are differences between the linguistic expression used and the message that is to be communicated. For this, and other similar reasons, it is vital that definitional issues be clarified prior to undertaking any scientific discussion. Thus, definitions of the key terms used in this study are provided for in the literature review to ensure clarity.

The focus of the present study was to explore airline pilots' perceptions of the advanced automated aircraft training climate after developing a valid measurement scale for the construct. The literature review set out in this chapter is useful in understanding, clarifying and defining the constituents of the training climate in the context of training to operate advanced civilian aircraft. Thereafter, the review examines some approaches learners may adopt when acquiring new knowledge. Measuring the psychological components of the training environment is discussed, particularly in respect of the complex aviation sector.

3.2 RESEARCH DELIMITATIONS

The study assesses a psychological climate in terms of its systemic association with advanced automated aircraft pilot training. A systemic organisational environment is described in terms of the external, internal and intermediate spheres of influence that apply in an organisational behaviour context (Leibold, Probst & Gibbert, 2005). The conceptual links emerge from a psychological bond between the organisation (in this case, the airline), the group (here, the instructor-trainee interface) and the individual (here, the trainee), as experienced by qualified advanced automated aircraft airline pilots. In this study, only the psychological processes (individual attributes) of the airline pilot (employee) were considered, whereas those of management were excluded from the scope of the study.

The unit of analysis in this study is therefore the perceptions of a particular group of airline pilots flying (operating or working with) advanced computerised (highly automated) commercial aircraft. These pilots typically occupy various positions at an airline organisation, where they command the aircraft (as captains), or are second-in-command (first officers) or, in some cases, third-in-command (second officers).

To investigate this training climate, the psychological theoretical streams were further limited to those dealing with perceptions, attitudes, and elements of the psychodynamics associated with personality and motivation; or associated with areas of learning, education and training. To gain an understanding of the human-machine system, only theories relevant to the operation of advanced automated aircraft were considered.

To simplify the presentation of definitions, tables were compiled to clarify some primary but complex aviation concepts used in this study. For instance, it was required to clarify early on, in Section 2.2 and 2.3, concise definitions and further explain that an advanced automated aircraft consists of two separate but related parts, namely the flight deck itself, and the airframe and associated mechanical sub-systems. It was necessary to initially differentiate two parts of the advanced aircraft because technological capability is based on the association between basic computerised external aircraft systems and computerised control and display units in the flight deck (Ausink & Marken, 2005; Dole, 1989). Overall, the pilot of an advanced aircraft would rely extensively on computer-based systems in order to control, monitor and manage the aircraft.

3.3 CLARIFICATION OF THE CONSTRUCT - TRAINING CLIMATE

According to Denison (1996), empirical research on phenomena associated with an organisational climate requires quantitative methods, because there is a need to generalise findings across social settings. It was therefore intended that the core components in the measurement scale developed in the course of this study would successfully assess perceptions of the advanced aircraft training climate of the airlines concerned. Hence, further elucidation of the term *training climate* is necessary at this point of the discussion in order to quantify and operationalize the final construct.

3.3.1 Introduction to climates

The literature differentiates between two separate climatic constructs, namely a *psychological* climate and an *organisational* climate (Chung, 1996:35; Denison, 1996:619). The psychological climate refers to making sense cognitively of the organisational environment. In this context, Denison (1996:621) has criticised authors who continually confuse “culture” with “climate”. Denison (1996) points out that an organisational climate refers to individuals’ subjective summated (average) sense made of interpersonal constructs, and their understanding of policies, procedures and structure. By contrast, Schein (2004) refers to the organisational culture as a set of group assumptions created after learning from a number of internal and external difficulties or problems. Climate researchers are “generally less concerned with [social] evolution but more concerned with the impact that organisational systems have on groups and individuals” (Denison, 1996:621).

It can then be established that employees (and in the present case, trainees) will adapt their behaviour based on their perception of the organisational climate (Fishbein & Ajzen, 2001; Tracey & Tews, 2005). For instance, according to Chung (1996), the behavioural characteristics of human beings are moulded by a plethora of simultaneous environments. After closer examination of this proposition, it can be deduced that the simultaneous environments in which human behaviour may occur are similar in nature to the business eco-system posited in Leibold, *et al*. (2005); that is, a business eco-system consists of systemically related environments (psychological, industrial and economic environments) and similar enterprises that fluctuate in unison within their corporate dimension. Therefore it may be conjectured that competiveness stems from the different business eco-systems and not specifically between similar enterprises within such eco-systems. One may conclude from such a proposition that the training of airline pilots by one organisation or company will be beneficial to the entire industry on the whole. However, the climatic construct developed for this research is associated with an individual enterprise (airline company) within the business eco-system. Future research into the perceptions of climates defined in terms of eco-systems may yield new information and possibly enhance industry safety.

For the purposes of the current study, the training climate consists of psychological bonding elements (based on multiple levels of internalised organisational influence, such as structure, management, leadership, corporate values and company strategic objectives). Katz and Khan (1966) argue that the mechanism of psychological bonding involves linkages between the human psyche and the organisational pattern. It has also been suggested that employees unconsciously seek out patterns within an organisation so as to bring about stability from available environmental information (Drucker, 1946). Such seminal arguments support the conclusion that the climate, in terms of the corporate training environment, has a significant psychological influence (exerted by the external environment), and thus exists within the mind of the beholder.

Because the core purpose of this research was to develop a perception measurement instrument of a particular climate, it was posited that the advanced aircraft training climate should consist of both organisational and psychological dimensions. However, this approach did not address the concern raised by Hellreigel and Slocum (1974:256) that there is confusion over whether the term “climate” refers to the attributes of people or the attributes of organisations. The current study adopts Denison’s (1996) view that it would be difficult to define a climate as referring *only* to people.

3.3.2 Airline training climate

A basic definition of the training climate is “all factors in the person, learning and organisation that influence [the] transfer of knowledge to the job function” (Rouiller & Goldstein, 1993:8). The airline training climate was conceptualised in terms of Rouiller and Goldstein’s definition, and two additional components were also included to create a holistic construct: firstly, a theoretical part (technical knowledge on the aircraft) and, secondly, a practical part (flight simulator and route training).

One of the characteristics that distinguish a training climate from a general business climate is the immense breadth and depth of interpersonal organisational behaviour relationships found in a training climate. The organisational behaviour standard model is entrenched in concepts originating from a broad spectrum of behavioural sciences; in particular, from psychology, at the microscopic or individual (person) level; from sociology, at a group or intermediate level; and from anthropology, at a macroscopic (organisational) level (Robbins *et al*., 2004).

Psychology refers to a scientific investigation into the mind of the individual, whereas the role of sociology in the behavioural sciences is to study people in relation to others (Argyris, 1957; Schein, 2004). An umbrella discipline covering the study of communities in the behavioural sciences is anthropology – "much of our current understanding of organisational culture, organisational environments and differences between national cultures, is the result of the work by anthropologists or those using their methodologies" (Robbins *et al*., 2004:10). Leibold *et al*. (2005) describe the interaction of organisational components at these three distinct levels as a kind of systemic choreography, and have classified this interaction as a business eco-system. Leibold *et al*. (2005) clearly demonstrate how the boundaries between human behaviour and business behaviour are somewhat blurred in theory, which implies that building an appropriate construct to measure human perceptions of organisational systems can be tricky.

Katz and Khan (1966) postulate that what they call *psychological bonding* closely links the mechanism of interaction between the layers of an organisational structure. Thus, it is reasonable to argue that a training climate may be described in terms of a business eco-system (Leibold *et al*., 2005), and can therefore be appropriately measured by a perception scale. The primary objective of the study can therefore be accomplished only by accepting the proposition that that perceptual or psychological bonding in human beings occurs when they interact with an organisational eco-system.

The construct of perceptions of the advanced automated aircraft training climate is then defined in terms of the aforementioned organisational behaviour concepts and of psychological bonding theory. It is relevant to associate the behaviour of employees in an organisation with business-related measures, because the current literature calls for an improvement in organisational behaviour research through the use of quantitative methods (Ellis, 2010). For instance, Tracey and Tews (2005:355) propose "strengthening" organisational research by operationalizing specific constructs, particularly with regard to training and education.

3.3.3 Climate constructs associated with the airline organisation

This part of the literature study critiques the paths taken by various researchers in conceptualising factors associated with a generic training climate construct and its influence on the airline organisation.

Leibold *et al*. (2005:55) argue strongly in favour of adopting systemic thinking when attempting to measure constructs of organisational behaviour because "the emphasis on the parts has been called mechanistic, reductionistic, or atomistic; the emphasis on the whole is termed holistic, organismic, or ecological. In modern science, the holistic perspective has become known as *systemic*, as opposed to systematic". This premise was adopted in building an airline-related climatic construct. The intention was thus to develop a systemic measurement construct.

When trainees have a good understanding of the systemic training environment, and this understanding is coupled with other moderating variables, such as previous experience, they are able to cope more effectively with the cognitive and physical demands of airline training (Davis, Fedor, Parsons & Herold, 2000). The self-efficacy gained in the individual, stemming from this closed loop environment (input-output-feedback), is demonstrated in the positive results of a training outcome (Davis *et al*., 2000).

Some other prominent researchers in the aviation training field argue for pursuing an interconnected solution to the problems of measuring perceptions related to an aviation training environment (Telfer & Moore, 1997). However, many of these models do not use contemporary systemic theory (Leibold *et al*., 2005). The aim was therefore to connect contemporary ideas of a training climate construct with that of aviation related human factors.

Table 8 contrasts some of the more important elements that emerged from seminal research. It is believed that such elements may be closely associated with the development of any new organisational training climates. Table 8 furthermore, shows that there are definite layers to any such climatic system, even where these layers are not explicitly commented on. These layers are interconnected and it is accepted that an interaction occurs which dictates the pattern of the perceived training climate. It should be noted from Table 8 that, terms such as *policies*, *supervisory, perceptions* and *individual characteristics* point to the three organisational behaviour dimensions – the macro, meso and micro levels.

Table 8: Chronological list of training climate elements

Chronological literary source	List of associated elements
Kozlowski and Hultz (1987:85)	1. Supervisory and trainer support. 2. Innovation policies. 3. Training job assignments.
James, Jones and Ashe (1990:110)	1. Individual and psychological attributes. 2. Trainee cognitive representation of practices and procedures. 3. Shared perceptions (trainee and trainer). 4. Situational cues that either inhibit or facilitate learning.
Rouiller and Goldstein (1993:45)	1. Consequences from learning or training. 2. Behavioural cues exhibited by supervisors, peers and subordinates. 3. Factors associated with the person, training and instructor and organisation.
Holton, Bates, Seyler and Carvalho (2000:68)	1. Contents of training programme. 2. Design of training programme. 3. Characteristics of the individual. 4. Motivational constructs of the individual. 5. Features associated with the working environment. 6. Preparation of training. 7. Outcomes of learning.
Tracey and Tews (2005:86)	1. Components of transference of learned knowledge and skills. 2. Components of shared and aggregated knowledge.

Tagiuri and Litwin (1968:32) define an organisational climate as "the relatively enduring quality of the total [organisational] environment that (a) is experienced by the occupants, (b) influences their behaviour, and (c) can be described in terms of the values of a particular set of characteristics (or elements) of the environment". This definition suggests that the training climate should in turn also consist of dimensional layers at an organisational, group and individual level. Therefore, the training climate associated with the advanced aircraft was constructed three-dimensionally (that is at a macro, meso and micro levels of analysis).

Hellreigel and Slocum (1974:256) refer to the organisational climate as "a set of attributes which can be perceived about a particular organization and/or its subsystems, and that may be induced from the way that organization and/or its subsystems deal with their members and environment". Hellreigel and Slocum (1974) therefore suggest that the way an employee *perceives* his or her environment should then be connected with a climatic construct. Similarly, Tracey and Tews (2005:355) describe the organisational climate as "a much broader, multidimensional perceptual variable", and point out, "specific dimensions or factor definitions should be determined by a specific criterion or criteria of interest". Therefore it may be expected that an airline pilot who experiences training to operate an advanced aircraft will develop some perception of the climate when meaning (pattern recognition) is associated with the learning that takes place (Neal, Griffin & Hart, 2000).

The goal of any structured learning process is to implement a range of observable behavioural changes in a trainee over a defined length of time. Cattell *et al.*'s (2002) model elegantly hypothesises a link between psychological factors, education and training. This model (see Figure 11) was adapted for the conception of the present research construct to measure an aviation training situation.

When one considers the complexity of a highly structured environment, combined with the interplay of a range of behavioural factors, it is clear that learning to operate an advanced aircraft can be a demanding task for an average person. Furthermore, the sheer complexity of advanced technology obviously requires an inherently structured learning process occurring over time. Because a vector is a function of time, that is, it maintains a relative position in space, one can argue, using Cattell *et* *al.*'s (2002) model; that an airline pilot's structured learning process for operating advanced aircraft includes measuring the rate of behavioural change produced by learning, mediated by factors such as ability, personality, motivation and personal states (see Figure 11). There is an abundance of similar models of the attributes of education and learning in a climatic context, but Catell *et al.*'s (2002) model was preferred in the current study because of the elements of psychology associated with it.

Figure 11: Mathematical relationship between structured learning and flight deck behaviour

1. $\bar{B} = f\left(\mathrm{A.P.M.S}\right)$
2. $\bar{B} = f\left(\mathrm{Time}\right)$

Structured airline training for advanced aircraft =

Rate of change in behaviour (B) as a function of:

Ability (A), Personality (P), Motivation (M), Personal State (S)

Source: Own derivation

Likert (1958) suggests that vector changes in behavioural domains similar to that described in Figure 11 allow behaviour to be measured in terms of interactive organisational elements as a function of time. Figure 11 mathematically depicts the variables associated with the demands of a regimented training programme. It also shows that, in a multidimensional system, core changes at the individual level of analysis are likely to occur.

Recently, airline training organisations have begun to identify and accept the critical role that individual factors can play in determining the outcomes of flight training (Abbott, 2010). They are responding to the conclusions of research that attributes as much as a third of aircraft accidents to enduring personality traits (Abbott, 1995; Helmreich, 1987, 2002; NTSB, 2009). In this regard, it is relevant that Figure 11 as derived from Cattell *et al.* (2002) suggests a behavioural prediction equation in line with the complexities associated with a structured learning model such as that which may be expected in a modern airline organisation. Cattell *et al.* (2002) further demonstrates that a structured learning process takes place across the aforementioned vector changes in terms of particular behavioural domains which have been incorporated within the model devised in Figure 11. Specifically, these personal factors tend to associate with intrapersonal psychological constructs such as:

- ability;
- personality;
- motivation; and
- personal state.

These domains were considered in the initial item pool generation and development of the final measurement construct (see Figure 13), so as to achieve the objectives of the current study.

It appears that individual factors play a fundamental role in airline organisations and that these factors in turn have a significant impact on the production of competent advanced aircraft pilots. The research model under investigation thus assumes that ripples in the training environment originate at an individual level to begin with.

3.3.4 Contextualising the advanced aircraft training climate

Social systems such as an organisation, lack the fixed demarcations commonly found in physical or biological systems, and instead consist of a combination of events that are inseparable from the functioning of the organisation (Katz & Khan, 1966). Thus, Katz and Kahn (1966) argue that organisations and societies behave in complex patterns and that the behaviour of each individual within an organisation is based largely on the requirements of the larger pattern. This implies that the outcome of an aviation training instruction effort is the result of the motivation and strategies adopted by the trainee pilot, the instructor’s value system, ability and knowledge, and by the very nature of the managerial pattern (Telfer & Moore, 1997:2). These elements are listed in Table 9. In addition, Table 9 and Figure 12 depict core elements of the advanced automated aircraft training climate, positioned as a systemic model. The structure of the model is largely based on the proposition that social systems exist within patterns of psychological bonding in terms of Katz and Khan's (1966) seminal proposal, as discussed earlier.

Table 9: Aviation-related psychological elements of a training climate

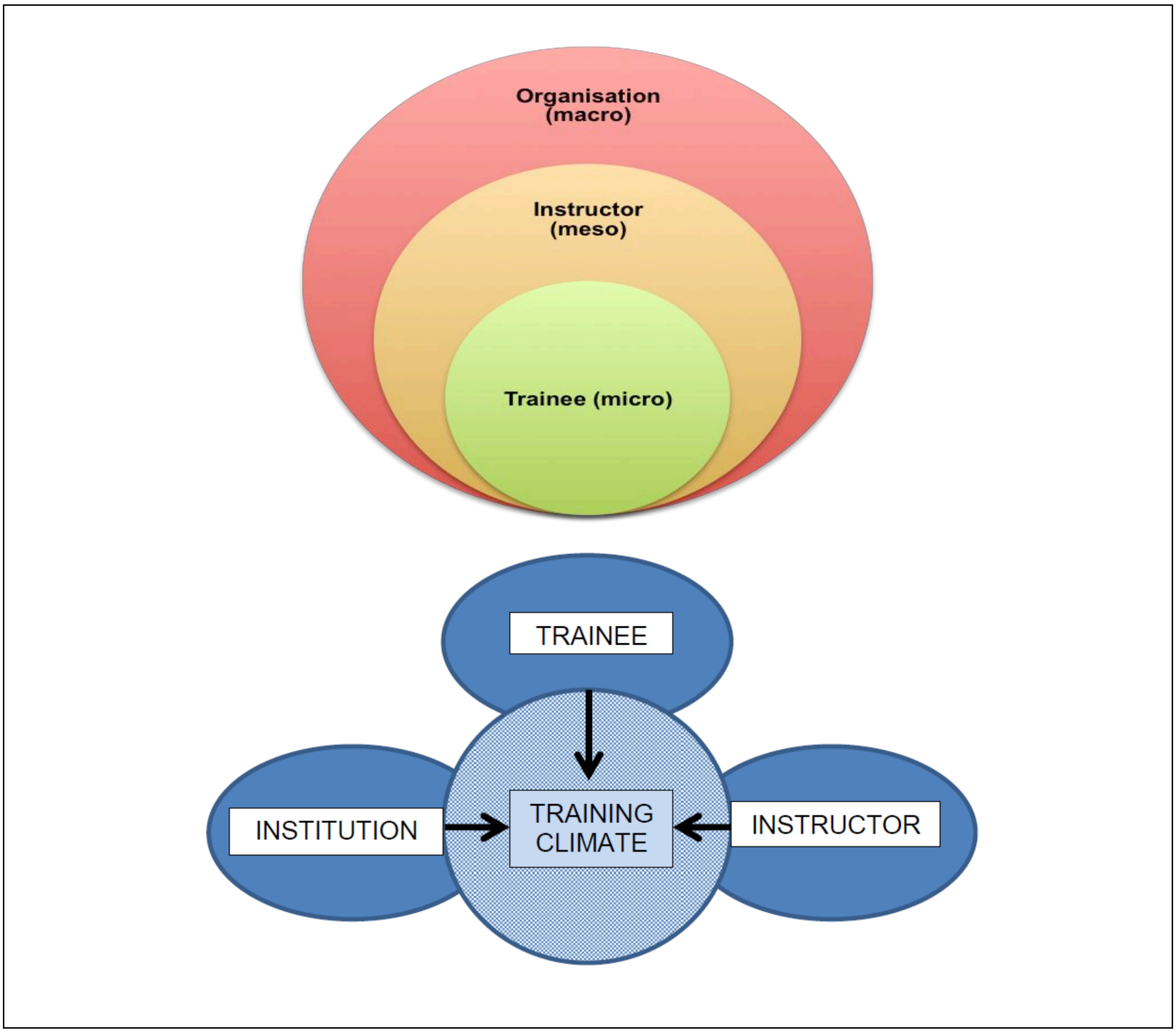
Trainee Pilot (micro sphere)	Perceptions of learning and psychological self (academic, social).
Instructional Group/ Instructor-Trainees (meso sphere)	Perceptions of teaching and interaction with instructor and co-trainee.
Airline Operator (macro sphere)	Perceptions of the organisational training atmosphere, structure, policies, standards and planning.

Source: Adapted from Katz and Khan (1966), Leibold *et al.* (2005), Roff (2005) and Telfer and Moore (1997)

The second diagram in Figure 12 depicts a shaded area envisioned as the airline training climate. In other words, a training climate is shown to be the intersection of three sub-constructs; namely, the institution, the trainee and the instructor or instructional dimension. Face validation of the concept is based on the effect presented in the variable overlap shown. The present research construct was then based on the premise that the *training climate* is an area common to all three independent variables (the trainee, the institution and the instructor). Although Figure 12 clearly shows where the individual instructor is positioned within the main construct, this dimension also refers interchangeably to the instructional domain consisting of the elements associated with an instructor-trainee relationship.

Figure 12 also illustrates and emphasises the importance of congruency between the various organisational levels. Congruency between levels assists the trainee, the instructor and the institution to attain the required degree of flight training success both efficiently and effectively (Tracey & Tews, 2005). According to Telfer and Moore (1997), a lack of congruence between levels will result in dissension amongst the various ranks, in turn weakening the system and contaminating the training climate. By analysing the current systemic situation through a measurement of output perceptions (held by qualified airline pilots), an organisation can attain sufficient knowledge to deal adequately with adverse training variables. Learner frustration and subsequent ineffective training may occur when departments operate in silos – in other words, when there is too little congruence between the goals and objectives in different sections of an organisation (Tracey & Tews, 2005). Moore, Po and Lehrer and Telfer (2001) emphasise the fundamental importance of aligning beliefs across all levels of an organisation, such as the management, teacher and student levels. Figure 12 clearly depicts the alignment and congruence attained within the main research construct. Alignment of independent belief systems is illustrated by the arrows, which emanate from each outer component (the trainee, the institution, and the instructor domain), pointing toward a core or centralised main component (the training climate). The integration of independent belief systems towards an aggregated pattern is summarised in the final hypothesised research construct (see Figure 13). The final measurement construct could only be developed after considering what approaches to learning trainees may adopt.

Figure 12: Representation of a systemic aviation training climate



Source: Adapted from Leibold *et al.* (2005:136), Roff (2005:15) and Telfer and Moore (1997:15)

3.4 APPROACHES TO LEARNING

Confucius (cited in Reynolds, 2001:174) once said: “Learning without thinking is useless; thinking without learning is dangerous.” Rosenow (2003) suggests that the art of thinking involves being mindful of the facts, by mitigating the impact of bias and prejudice with which knowledge is often peppered. An inability to learn is often coupled with an inability to think logically (Aronson, 1991). Most of the reasons generally associated with learning that is hampered are related to personal issues (see also Figure 11). Individual-based problems such as stress, low self-esteem, anger, illness, sleep deprivation, fatigue, laziness, depression and other effects have been cited as significant contributors to dysfunctional learning (Rosenow, 2003:47). This implies that factors originating from an individual or personal level of analysis are an important consideration when one needs to understand how human beings learn in terms of their perception of climate.

Some studies have objectively demonstrated that personality traits play a statistically significant role in influencing each individual’s character and behaviour, and thus the person’s approach to learning (Aronson, 1991; Cattell, 1946; Cattell *et al.*, 2002). From an aviation perspective, some research has shown that approximately one third of aircraft disasters may be directly attributable to individual enduring personality traits (Abbott, 1995; Helmreich, 1987, 2002; NTSB, 2009). Therefore, intuitively, aviation research specifically in terms of training should begin at an individual level. Furthermore Biggs (1987) posits that methods generally adopted by pilots for acquiring specific knowledge and skills to operate aircraft could differ significantly from person to person based on attitude.

Various psychometric scales have been developed and constructed to measure an individual’s capacity for learning, but most of these inventories are based on generic areas of educational research; such as at schooling and tertiary institutions. Adaptations of early education psychology theory, particularly from theory first formulated in the seminal Swedish studies conducted during the 1970s, have led to some of the modern versions of learning inventories and instruments. According to Biggs (1987), Biggs and Moore (1993), Entwistle and McClune (2004) and Ramsden (1992), learners tend to adopt one of two fundamental strategies, namely a surface learning strategy or a deep learning strategy. These broad strategies are characterised as follows:

- *Surface learners*: These students reproduce knowledge through rote memorisation. This phenomenon is further defined as the acceptance of new facts and ideas without much scrutiny (little thinking). Information is acquired at face value and stored in isolation with little or no connectivity. Students who use this strategy may often fail to distinguish principles from examples. A lack of interest in the content and the subject itself and a focus on merely obtaining a qualification characterises surface learners.
- *Deep learners*: These students gain meaning and understanding from knowledge based on a deeper level of acquisition (thinking). By examining additional facts or new ideas critically, deep learners can integrate new knowledge into existing cognitive structures. Focusing on the concepts that are needed to argue through a problem characterises an insightful level of knowledge acquisition.

Lasting individual personality traits, combined with environmental perceptions, can radically influence modern airline pilots' training and learning strategies. Warren (2004) suggests that organisations that encourage a deeper level of understanding should also ensure that subjects apply what was learned effectively. A deep comprehension of concepts is fundamental to integrating early understanding with present application, especially in a technical environment such as engineering, medicine and, in this case, operating advanced automated aircraft (Sherman, 1997; Wilson & Weston, 1989).

The importance of reflective learning is particularly important in an aviation safety context. This requirement was highlighted when airline pilots' training came under intense scrutiny during an air crash investigation (NTSB, 2009) after a Colgan Air Dash-8, Flight 3407, stalled while the airplane was on a final approach, and crashed into a suburb (Pasztor, 2009). The post-accident analysis revealed no mechanical malfunction in the aircraft itself (NTSB, 2009). However, the investigators found that the flight crew lacked a critical and fundamental understanding (trained knowledge) of their aircraft systems. The crew also misunderstood aerodynamic principles and there was substandard crew resource management (CRM), a fatal combination that led to the disaster.

More recently, analysis of the flight data recorders in an Air France Airbus A330 crash in 2009 led the manufacturer to advise pilots about the recommended abnormal and emergency procedures to be followed in case of an unreliable airspeed indication. In other words, Airbus appealed to pilots to reflect on basic aerodynamics and the effects of thrust setting versus actual airplane trajectory, something pilots are generally taught very early in their careers on basic aircraft. “Clearly it pointed to the possibility that mismanaging the plane’s speed had been one step in a cascade of on-board failures, leading to the crash” (Wald, 2009:11). By contrast, high levels of experience, coupled with highly effective training methodologies, averted disaster in a Qantas Airbus A380 incident, which occurred over Western Indonesia (Milmo & Webb, 2010). After severe damage to the aircraft’s number two engine, many fundamental flight systems were cut, resulting in spurious warnings emanating from the computerised monitoring system. The crew of this particular aircraft understood their equipment and the functioning of its related systems very well. Moreover, three of the four crewmembers were instructors for that aircraft type. Through reflective and effective reasoning, the team were able to determine which warnings were authentic, and which warnings were inconsistent with the problem at hand. This level of thinking and reflection enabled the crew to bring the aircraft safely back to the departure airport without any casualties.

The limited availability of academic information in the area of the application of learned theory to effective airline flight operations and supportive flight deck behaviour constitutes a knowledge gap in what is currently understood about the phenomenon. This has made it difficult for the current research to determine what exactly constitutes aviation training antecedents. Lowy (2009) has determined that there may have been some efforts to bridge this knowledge gap by scientifically analysing aircraft accidents and incidents, and also successful outcomes in adverse operational situations. As more data is gathered, it is becoming clear that understanding the role which flight training plays can guide solutions to future effective aircraft operation (Abbott, 1995).

3.4.1 Application of learning in the airline environment

In air crash investigation reports, inadequate transfer of theoretical or learned knowledge (which can only come from a deep conceptual understanding) to practical applications (actually flying an aircraft) is usually castigated as a sign of a lack of pilot *experience* (Pohlman & Fletcher, 1999). Such conclusions appear vague, and fail to pinpoint a root cause. For instance, a review of the most recent aircraft accident statistics reveals that many commuter or smaller airline operator accidents could be attributed to insufficient pilot experience levels (Pasztor, 2009; Patrick, 2002). However, research has also determined that some highly experienced airline pilots flying for large carriers also add to accident statistics. In many cases involving experienced pilots, it appears that insufficient experience in operating complex technology is a major contributor to accidents (Sherman, 1997). By contrast, Howell and Fleishman (1982) earlier pointed out that sufficient training and understanding in complex technology may mitigate or compensate for low levels of experience. Therefore, when examining human factor based aircraft accidents and incidents, a lack of systems understanding rather than actual experience levels should be considered a root cause. Paradoxically, some authorities have found that in an effort to contain costs, many airline operators cut back on some important aspects of pilot training. This has led the FAA in the United States to adopt far stricter oversight policies when it comes to pilot education, flight training and licencing (Moses & Savage, 1989), an example, which is being followed closely by their South African counterpart, the South African Civil Aviation Authority (2011b).

Telfer and Moore (1997) found that pilots exhibit competitive tendencies during training. Biggs's (1999) findings confirm this and suggest that the so-called *achieving* aspect of learning is strong among airline pilots. This construct describes those students who are enthusiastic about doing well in tests and exams (a deep sense of learning) and are competitive. In other words, these candidates are able to apply deep learned theory effectively, and then apply it in a competitive sense, because of their need for achievement. The need to achieve is strongly stimulated by a passion for the profession (Vermeulen, 2009). Thus, the *best pilots* are generally the ones who display a deep-seated passion for their chosen career path.

Moon (2004) contends that the premise that learning takes place in only two distinct dimensions – either in the form of superficial understanding (surface learning) or by truly understanding (deep learning) – may be a somewhat simplistic view of the situation. Moon's (2004) suggestion influenced the construction of the model used in the current research, because Moon demonstrates the existence of *multiple factors* responsible for affecting the approaches adopted by learners. These factors are:

- *Conception* – how a trainee envisions the learning process affects the method or the learning strategy that the trainee adopts.
- *Instructor influence* – teaching and assessment requirements determine the difficulty of the exercise and therefore have an impact on the learning approaches adopted by a trainee.
- *Demands* – a trainee's perception of the level of stress imposed by the process to meet the requirements to succeed influences the trainee's choice of learning approach.
- *Personal factors* – individual aims, goals and outcomes directly influence learning strategies.
- *Experience* – Moon (2004) is of the opinion that a trainee's prior knowledge of a subject, or experience with the subject matter, has an impact on the trainee's decision to adopt certain learning strategies. Familiarity with the current physical and psychological environment plays an equally important role.
- *Self-management* – a trainee's maturity determines the person's emotional orientation in terms of self-management and so has a direct impact on learning.

The proposed domains explicitly categorise the levels of learning strategically (Moon, 2004). It can be inferred that these domains are highly influential in channelling a trainee to adopt one approach rather than another, or in addition to another, for instance, adopting learning strategies with simple surface characteristics, or a strategy that promotes a fundamental and deeper understanding of the subject.

Lawshe's (1975) findings suggest that the psychological domain associated with training is another important factor to consider in understanding what learning strategies a subject may adopt. Higher order psychological variables are involved in learning, such as deductive and inductive reasoning. Moreover, the behavioural nature of the domain in question is characterised by the following factors or attributes (Lawshe, 1975:565):

- directly observable behaviour;
- reportable behaviour; and
- abstract behaviour.

Regulators recognise the various behavioural dimensions of training; for example, it is a subtle factor influencing the requirements for demonstrating competence during tests, evaluation or assessment. According to the Civil Aviation Authority (2011), the civil aviation technical standards (CATs) and the civil aviation regulations (CARs) require that, in order for an applicant to be deemed competent by a flight testing instructor, he or she must have demonstrated both knowledge of aircraft specific technical theory (abstract behaviour) and competent manipulation of the aircraft flight controls (directly observable behaviour). Both components are behaviours that are reportable and can therefore be documented; on the basis of such reports, an appropriate licence can be issued or denied. Lawshe's (1975) model is highly relevant in understanding both the psychological and behavioural aspects of learning.

3.4.2 Literature on airline pilots' learning styles and subsequent organisational impact

McManus, Keeling and Paice (2004) suggest that the ability to complete a job, workplace climate, stress and burnout are very closely linked to an individual's learning style. Personality traits can also play a significant role in the style that a person adopts (Cattell *et al.*, 2002). Personality elements are significant in predicting training success (McManus *et al.*, 2004). Aspects of personality such as conscientiousness and agreeableness are traits found in those pilots who are more likely to succeed in training (Berliner, 2006). Airline pilot recruitment specialists have capitalised on these theories by looking for specific personality attributes in the individuals whom they eventually hire (Telfer & Moore, 1997).

The early work of Kolb (1976) suggests that some conscious efforts can affect the learning style that a trainee adopts and subsequently have an impact on the success of the outcome of the application of the learned theory. Hence, Kolb's (1976) model presents two orthogonal dimensions (McManus *et al.*, 2004). An orthogonal model is plausible because it combines bipolar dimensions of cognitive growth based on the transformation of experience (from both active behaviour and abstract behaviour). Obtaining a pilot's licence thus requires mastering both orthogonal dimensions (SACAA, 2009). Therefore, the hypothesised measurement construct for the present study was designed by considering orthogonal levels of learning. Items concerning an individual's personal attributes were also considered important for inclusion in the operationalization of the research model.

Because both motivation and personal or individual attributes of all active participants in a training environment (instructors, trainees) play increasingly important antecedent roles in successful or unsuccessful learning styles, broadly structuring learning into a surface or deep process can be problematic in an airline training organisation (Moore *et al.*, 1997). Ashcroft and Foreman-Peck (1994:22) posit that there is a basic challenge in categorising learning styles because "it is easy to induce a surface, reproductive approach by structuring the learning demand, but very difficult to induce a deep approach...it may be possible to promote a deep approach by altering the learning demands for some, but not all students". Deep approaches to learning are promoted and cannot be built into a structure, because this approach stems from an individual level.

Depending on circumstances, in certain instances, a trainee may indeed find the need to adopt a surface learning style, because in aircraft training it may be necessary to memorise specific material (Airbus, 2011b), as some aspects of piloting do require rote memorisation. So, for example, rote memorisation of certain checklists, call-outs and standard operating procedures (SOPs) *verbatim* are an important part of training. This technique may be analogous to the way an actor memorises the lines of a script. For example, a mechanical procedure is required when a landing is aborted. The pilot is required to verbalise and manipulate the aircraft in a precise sequence so as to stabilise the aircraft's trajectory after such a manoeuvre.

Rote memorisation of procedures is limited mainly to aspects of flying that require immediate and deliberate action from the pilot. This is why the basic philosophy of advanced aircraft manufacturers is based on the premise that pilots should have as few memorised action items to deal with as possible. Airbus, for instance, requires a pilot to perform manoeuvres by memory for only seven non-normal operations (Airbus, 2011a) in an automated aircraft, in contrast to the operation of older analogue-type large commercial jets, which required a pilot to memorise almost three times that number of items (Degani *et al.*, 1995). This discussion therefore suggests that pilots who are successful in training have the ability both to memorise and to understand topics related to flying.

The difficulty in finding good theory related to airline pilots' learning styles and approaches is similar to the complexity of probing learning in other technical professions, such as the learning styles adopted by medical doctor trainees. These two groups of learners share some similarities: a study conducted by Wilson and Weston (1989) found that the working hours of junior anaesthetic doctors were comparable to those of airline pilots. Both jobs include intense episodes that make high physical and cognitive demands, coupled with significant periods that involve simple monitoring. McManus *et al.* (2004) refined the categorisation of so-called surface learners after analysing medical students undergoing training. Classing the surface learners as either surface-disorganised or surface-rational subcategorised the surface or rote-learning group. A surface-disorganised student is predicted by factors relating to high scores on neuroticism (emotionally unstable) and lower levels of conscientiousness (carelessness or haphazard traits), when described in terms of the Big-5 personality models.

Although a number of contemporary psychological theories describe personality models, the theories are mainly derived from an earlier Big-5 model of personality. According to Cattell (1946) and Costa and McCrae (1992), five significant factors can account for the majority of the variability in all personality types. Generally, these factors with underlying correlated trait variables were labelled as follows (Cattell, 1946; Costa & McCrae, 1992):

- *openness to experiences*, with a trait scale ranging from consistent or cautious at the lower end, to inventive or curious at the upper end;
- *conscientiousness*, with a trait scale ranging from over-easiness or carelessness at the lower end to efficiency and being organised at the upper end;
- *extroversion*, with a trait scale ranging from solitary or reserved at the lower end, to extroverted, outgoing or energetic at the upper end;
- *agreeableness*, with a trait scale ranging from unkind, aloof or cold at the lower end, to compassionate and friendly at the upper end; and
- *neuroticism*, with a trait scale ranging from emotionally unstable, sensitive and nervous at the lower end, to emotionally stable, secure and confident at the upper end of the scale.

The general personality models can be linked to various assumptions of learning. For instance, Rogers and Skinner (1956) propose that learning essentially takes place in three dimensions. Firstly, learning can be observed as a change in behaviour. Similarly, a personality type is manifested in overt behaviour. For example, a person displaying low levels of conscientiousness may appear disorganised and haphazard. It can be assumed that these characteristics are an indication of inherent traits within the individual (Cattell, 1946). Secondly, the environment is an important antecedent of education and thus shapes learning. This implies that the training climate plays a significant role in learned operant behaviourism. Thirdly, the contiguity of events determines the level of conditioning and ultimately the success of the learning. Therefore, personality types that are strong on the conscientiousness scale are more consistent or controlled; and their level of organisation results in stronger bonds being forged, that is, in better conditioning (Costa & McCrae, 1992). In terms of the learning strategies possibly adopted by trainee pilots, surface-rational students are characterised by strategic learning during their early years of study. The early learning experience is generally coupled with far less openness to experiences (a more cautious approach); however, later on, it is linked to higher levels of conscientiousness. During the initial stages of learning, this kind of trainee attempts to gain as much knowledge as possible using minimum cognitive effort, with the aim of passing exams and tests (a surface or rote memorisation method is then usually adopted). After obtaining specific qualifications and reaching a level characterised by stability and less stress, these trainees may then focus on gaining a deeper understanding of their subject. Conscientious trainees are able to link the reward of receiving praise and admiration for their knowledge with many other reinforcements, which can greatly increase the likelihood that the trainee will continue to behave in this manner on subsequent training courses. According to Rogers and Skinner (1956), the operation of a learner's behaviour within the environment is shaped by reward and punishment. In addition, the changes in a trainee's behaviour are measurable in terms of both successes and failures in actual flying and written examinations. In part of a study that analysed new pilots undergoing conversion training at a large airline company, Telfer *et al*. (1996) found that these participants scored above the mean for the *achieving* factor than was the societal norm. Although it was concluded that airline pilots appeared to be very competitive in their chosen learning strategies, they did seem to use a substantial amount of rote (surface) methods during the early stages of their training (Telfer *et al*., 1996), very similar to the medical students mentioned by McManus *et al*. (2004). This appears to be in line with the theory that students strategically use the memorisation of material to pass specific assessment levels during initial training. Regrettably, in analysing aviation learning strategies, Telfer *et al*. (1996) did not go any further in defining or splicing the surface dimension in their particular study, unlike the study on the medical students conducted by McManus *et al*. (2004). Therefore, the results are inconclusive regarding why pilots adopt a surface learning technique. Furthermore, the comparison of the medical students group with the pilot trainee group is hypothetical. No quantitative statistical comparative analysis has been conducted to verify the significance of the hypothesis. Nonetheless, the data provided by the research on the medical students served as a guide in developing the present research construct. Most of the literature on learning approaches favours a dichotomous view and does not offer a formal definition, but rather describes the nature and attributes of learning styles within specific organisational contexts. The role played by the organisation has caught the attention of educational researchers exploring the factors that influence students' learning approaches (Schaap, 2000). In an organisational behaviour context, authors prefer to describe the interaction in the learning environment according to the elements of the macro, meso and micro spheres (Ashcroft & Foreman-Peck, 1994; Berliner, 2006). Table 10 lists some of the elements associated with such levels of analysis.

Table 10: A synthesis of the elements affecting learning at different levels of analysis

	ELEMENT		
Source	Airline Operator (macro sphere)	Instructor-trainee group (meso sphere)	Student (micro sphere)
Ashcroft and Foreman-Peck (1994)	Bureaucratic requirements. The type of assessment systems in place. Nature of the environment, either caring or distanced.	The approachability of the instructor. Teaching style and method. Conact with the trainee. Size of the group being taught.	Style of the student. Personal goals and values. Social expectations. Interest in the subject being taught. Emotional well-being.
Berliner (2006)	Socio-cultural history. Social class. Status of the organisation.	Level of ‘parental’ education offered. Characteristics of the instructor or teacher. Intelligence of the instructor. Ethnicity of the instructor. Sociability of the instructor. Interactions between instructor and student. Media type used for instruction.	Intellectual ability of the student. Personal values. Motivation to learn the subject being taught.

3.5 MEASURING LEARNING

“The work of education is to make changes in human minds and bodies. To control these changes we need knowledge of the causes which bring them to pass” (Thorndike, 2007:3). Early attempts to describe how human beings approach learning and their eventual methods to acquire the skills to complete complex tasks typically focused on why people selected dichotomous cognitive strategies (Moon, 2004). As already indicted in Section 3.4, according to Ashcroft and Foreman-Peck (1994), early research demonstrated two ways in which adults approached learning, namely *surface learning* and *deep learning*. These two concepts are still often the basis for measurement when researchers attempt to measure perceptions of learning.

The evolution of educational research has produced many inventories to explain and measure how learners acquire knowledge (Biggs, 1987; Moon, 2004; Pololi & Price, 2000), for instance, the Study Process Questionnaire (SPQ) developed by Biggs (1987). These inventories have been adapted by researchers, and in most cases, have been used in the measurement of college and medical students’ approaches to learning. The literature reveals that students’ and adults perceptions of their learning environments have been diagnosed with the use of several complementary and conflicting inventories over the past decade, from the primary through to the tertiary education levels. However, a review of the available resources also shows that accurately measuring airline pilots’ perceptions would require the development of a new inventory.

Both qualitative (Denzin & Lincoln, 2005a, 2005b; Seabrook, 2004) and a multitude of quantitative methodologies (Schaap, 2000) have been used to analyse the educational environment. However, an extensive search of leading academic electronic databases (EBSCOHost, Emerald, Google Scholar, Proquest, ScienceDirect, Informaworld) and various hardcopy library resources suggested that very limited information is available regarding research in the last five years into learning measurements related to the educational environments of trainees in the aviation industry. Furthermore, within the last two decades, very few researchers have attempted to analyse, describe and publish results of an analysis of the learning environments of airline pilots, particularly those operating advanced aircraft (Sherman, 1997; Smith & Dismukes, 2000; Telfer & Moore, 1997). Most adult learning measures reviewed were developed in the medical and tertiary education fields – little attention has been paid to the perceptions of airline pilots operating highly advanced equipment. Nevertheless, the current study relied on models conceptualised from earlier generic learning research results. These were used as a platform for developing a measure to assess airline pilots' perceptions of the training climate associated with their training on advanced aircraft. The domains of previous measures, from other industries, were also part of the framework used to determine what constitutes a learning environment and thus a training climate. Based on this tentative framework, appropriate items were then generated to construct an assessment questionnaire used to assess content validity of the main construct (see Tables 19 to 21). The learning measures consulted for the construction of a working framework, which guided the present study, are summarised in Table 11.

Table 11: A chronological synthesis of some important learning inventories

Source	Instrument used for reference
Rothman (1970)	Medical School Environment Index
Levy (1973)	Learning Environment Questionnaire
Kolb (1976)	Learning Style Inventory
Marshall (1978)	Medical Schools Learning Environment Survey
Myers and Briggs (1979)	Myers-Briggs Type Indicator
Honey and Mumford (1982)	Learning Styles Questionnaire
Biggs (1987)	Study Process Questionnaire
Moore, Lehrer and Telfer (1997)	Pilot Learning Processes Questionnaire
Sherman (1997)	University of Texas Aviation Automation Survey
Entwistle (1998)	Approaches to Study Inventory
Pololi and Price (2000)	Learning Environment Survey
Schaap (2000)	Learning Approaches Questionnaire
Roff (2005)	Dundee Ready Education Environment Measure

Table 11 lists some of the learning, training and educational inventories established during the last three decades. These inventories were appropriately consulted in the development of the present scale. It can be observed that only two of these instruments diagnosed aviation-related educational environments *per se*, namely the Pilot Learning Processes Questionnaire developed by Moore *et al.* (1997) and Sherman's (1997) University of Texas Aviation Automation Survey. Only Sherman's (1997) survey contained an analysis of advanced flight deck automation training. However, Sherman's (1997) study did not address the learning environment or training climate *per se,* leaving many questions unanswered. Moore *et al.*'s (1997) Pilot Learning Processes Questionnaire explored the perceptions of pilots who operated older generation analogue-type aircraft. The study by Moore *et al.* (1997) does, however, provide a systems framework for the analysis of airline pilots' organisational behaviour, aspects of which were used in the framework to develop the theoretical model proposed in the current study. A review of these inventories thus revealed the need for new research. Additionally, combining the ideas from the two aviation-related surveys and the generic learning inventories developed over the last thirty years provided some constructive guidance for the construction of the new measurement tool developed in this study.

3.6 HYPOTHESISING AN EXPLANATORY MODEL OF THE RESEARCH CONSTRUCT

The alignment of beliefs (organisational pattern generation) in an advanced automated aircraft training environment ensures that a well-maintained system is in place to generate competent pilots during transition training (Bent, 1996). The challenge of measuring trainees' perceptions of this system provided the seed for the development of a hypothetical measurement model. By synthesising the information gleaned from the scholarly literature review, a conceptual theoretical model to measure perceptions of the advanced automated aircraft training climate was constructed. Figure 13 depicts the core theoretical model used in the operationalization of the construct and development of the final measurement scale. The goal of this part of the research was to measure the extent to which the theoretical model provides adequate coverage of the investigative objectives and propositions. Based on the mechanisms for psychological bonding (Cattell *et al.,* 2002; Katz & Khan, 1966), the research model was also based on three levels of fundamental analysis:

- the organisation;
- the group; and
- the individual.

Throughout the literature, a so-called open systems concept was a common finding in describing successful aviation training organisations (Andrews & Thurman, 2000). Adopting a systemic approach also allowed the final research construct to use seminal theory as a conduit (in other words, arriving at the unknown from the known). A total of 17 critical measurement domains, encapsulated in three fundamental dimensions, were ultimately hypothesised for the research theoretical construct. These three dimensions are the following:

- Dimension 1: The micro sphere, derived from *psychology* – the measurement domains were
    - Learning for technology (Le);
    - Motivation to train (Mo);
    - Personality (Per);
    - Training Stress (Sts); and
    - Training decision-making (Dm).
- Dimension 2: The meso sphere, derived from *sociology* – the measurement domains were
    - Training group dynamics (Gd);
    - Intergroup training behaviour (InGB);
    - Simulator training teams (Ste);
    - Training conflict (Co);
    - Power (Pr); and
    - Communication (Com).
- Dimension 3: The macro sphere, derived from *anthropology* – the measurement domains were
    - Training culture (Cu);
    - Knowledge Environment (En);
    - Structure (Str);

- Training Policy (TPol);
- Training standards (Std); and
- Training Planning (TPla).

Table 12 shows a list of the various psychological, behavioural and learning theories that were consulted in developing an operationalization method for the model (see Figure 13). The table listing the theories provides only an overview of the vast amount of work done in the field (an in-depth discussion of these theories was beyond the scope of the current thesis and therefore specific references are not mentioned). In addition, the theories that were assessed here are well known as seminal research in psychology as an academic discipline; such as those of Maslow, Vroom and Schein. In addition an in-depth analysis of these theories at a doctoral level would have been superfluous. The conjectural item pool that had received content validation from a panel of experts (discussed in Chapter 4) was originally created by adapting the root theories listed, and also categorising them at three specific levels of analysis (see Table 12). According to Corsini (2002) in Roeckelein (2006:X), "[a] *theory* is a body of interrelated principles and hypotheses that purport to explain or predict a group of phenomena that have been verified largely by facts or data; *hypothesis* is defined as a testable proposition based on theory, stating an expected empirical outcome that results from specific observable conditions".

The present study's hypothetical concepts and constructs were derived from seminal theory. Additionally, the final construction of an integrated hypothesised main research construct (see Figure 13) was based on the relationships that are believed to exist between these seminal theories from an aviation industry perspective. The model forms the foundation of the current research and depicts the main construct of measurement, namely *Perceptions of the Advanced Automated Aircraft Training Climate*. The rationale of the hypothesised research model stems from fundamental principles found in the organisational behavioural sciences (that is, at three systemic levels of analysis).

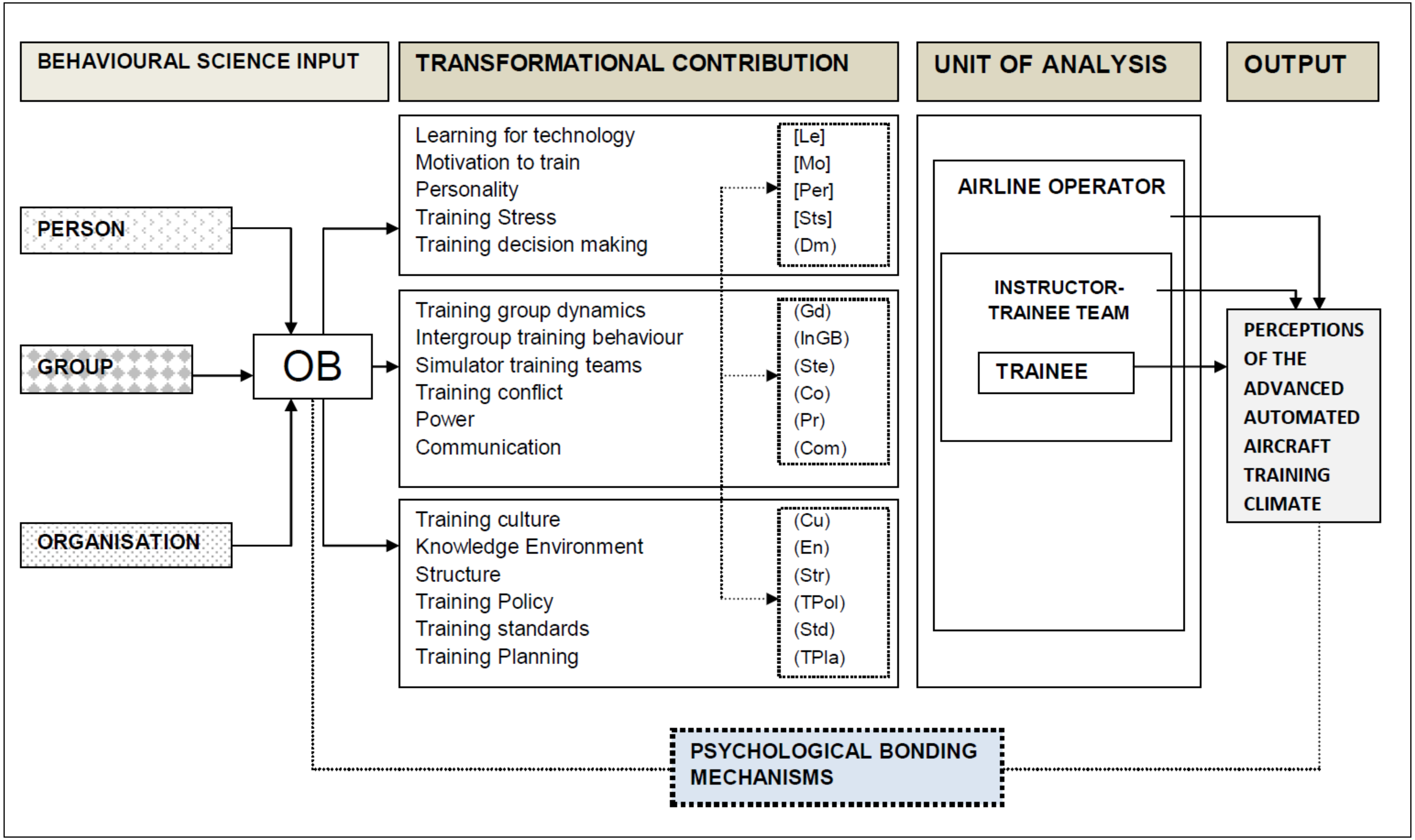
Figure 13 clearly shows that the research construct is multidimensional in nature. According to Cooper and Schindler (2003), multidimensional constructs consist of simpler and more concrete concepts (in this case, 17 such concepts). Such a conceptual model follows from the work supported by Cattell's (1946) theories of mathematical behavioural prediction modelling (also see Figure 11) that show behavioural-learning outcomes or gains, which are modelled from formulae determining the combined effects of intrapersonal psychological variables.

Table 12: Root theories considered in the construction of the theoretical model

Micro sphere	Meso sphere	Macro sphere
Alderfer's (Existence, Relatedness and Growth) ERG theory	McClelland's acquired needs theory	House's path goal theory
Maslow's hierarchy of needs theory	Adam's equity theory	Evan's theory of leadership
Friedman's theory of type A and type B personalities	Vroom's expectancy theory	Schein's theory of culture
Rotter's locus of control theory	Taylor's theory of group dynamics	Deal and Kennedy's organisational behaviour theory of culture
Bandura's social learning theory	West's theory of teams	
Ajzen and Fishbein's theory of reasoned action	Homan's group dynamics theory	
Ajzen's theory of planned behaviour	Brigg and Meyer's group decision making theory	
Roger's and Maslow's humanistic theory	Janis and McCauley's theory of group cohesion	
Cattell's state trait theory		
Allport and Cattell's enduring traits theory		

Source: Adapted from Desler, (2002); Furnham (2008); Roeckelein (2006)

Figure 13: Hypothesised model of the main research construct



Source: Author

3.7 CONCLUSION

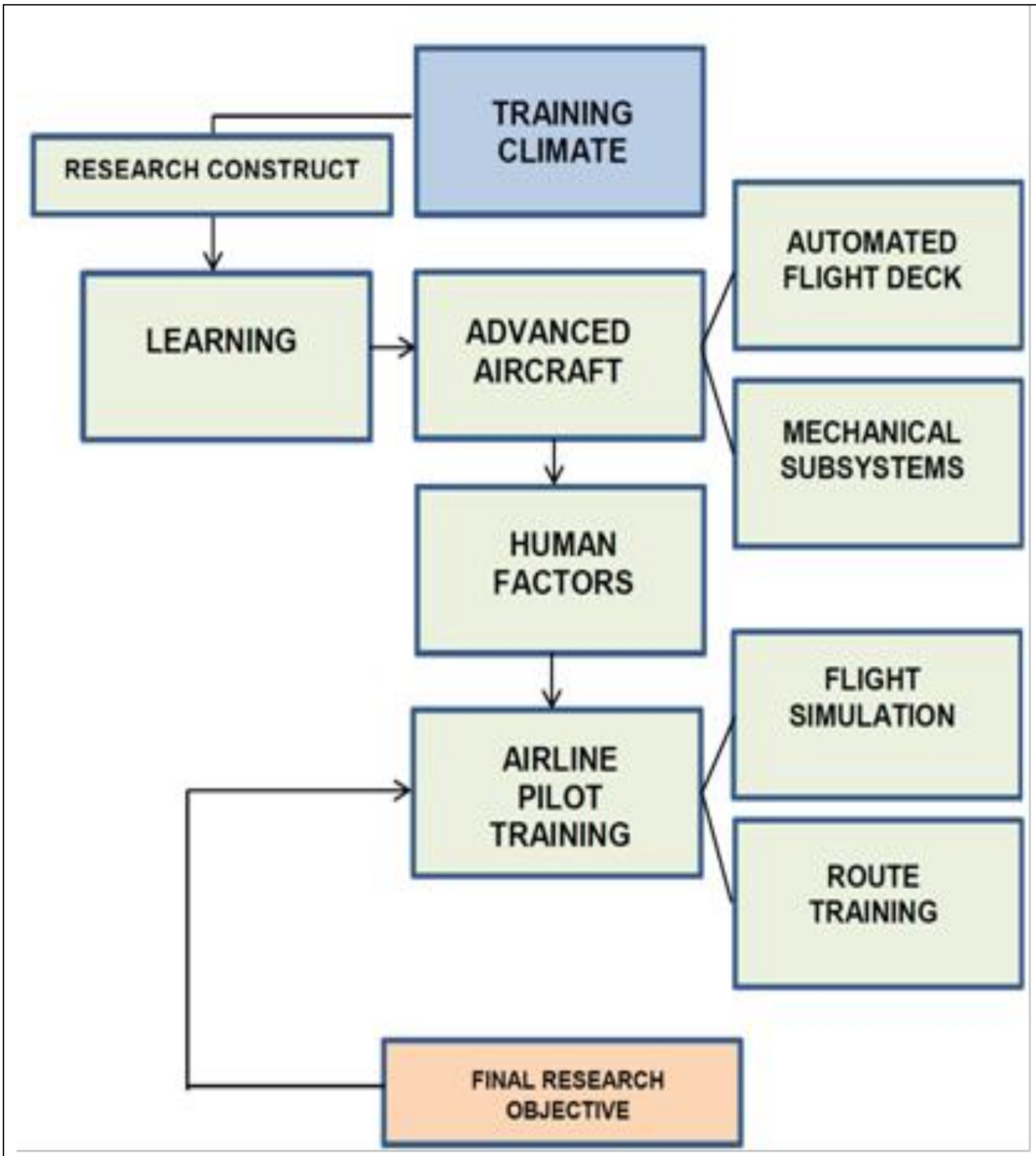
This chapter has focused on education and training, on learning and on the environment in which such learning takes place in an airline training context. The construct of the training climate was discussed, because the training climate played a fundamental role in the design and final construction of the hypothetical measurement model. Theory from the behavioural, social and psychological sciences was examined as antecedents to the current study's objectives. Concepts were borrowed from generic organisational theory to allow for an understanding of learning in an aviation paradigm.

The importance of understanding how learning takes place, and specifically the current and earlier measurement methods adopted by various scholars, were also discussed. This provided a logical build-up to the important literature reviewed in developing the current research's operational model, where the employee refers the pilot at an individual level, while the team refers to the instructor, co-trainee-trainee pilot domain, at a group level; and finally, at an organisational level, the organisation referring to the airline.

Modern airline organisations can only reap the rewards of longer-term safety spin-offs emanating from a clearer understanding of learning with regard to advanced technology aircraft if knowledge of pilots' perceptions regarding the training climate becomes an integral part of their everyday management practices. This will require the adaptation at virtually all levels of the enterprise to accommodate the challenges of the human-advanced-machine learning environment.

Figure 14 presents an integrated summary of the focus of the literature reviewed in Chapters 2 and 3, linking the theory to the design and development of the final hypothetical measurement construct, and leading to the ultimate research objective.

Figure 14: Summary of the focus of the literature review and its integration with the research objective



In the next chapter, the research design and methodology adopted to attain and analyse the relevant data are defended and discussed.

CHAPTER FOUR:

RESEARCH AND STATISTICAL METHODOLOGY

4.1 INTRODUCTION

> *If we knew what it was we were doing, it would not be called research, would it?*
>
> (Albert Einstein, cited in Calaprice, 2005:15)

The discussion thus far has dealt with gaining a deeper understanding of the formulation of a hypothetical model of the measurement construct, namely *perceptions of the advanced automated aircraft training climate.* The literature review built a foundation for the model and the subsequent framework, which demonstrated *why* it is important to develop such a measurement. This chapter contains a discussion of *how* the research was conducted. The discussion includes the type of research design employed, the philosophy that underpinned the construction of the measuring instrument and a rationale for the methodological approach adopted in this research.

According to Babbie (2010), possibly the most effective methodology that can be used to gather information about a large population, is the survey method. Surveying an appropriately selected sampling frame of the population is particularly useful when a researcher wishes to measure attitudes and orientations towards hypothesised phenomena (Schreiner, 2010). Furthermore, it is suggested that the survey method is an ideal vehicle for the purposes of conducting descriptive, exploratory and explanatory studies (Babbie, 2010; Cooper & Schindler, 2003).

When designing the present study, heed was taken of Pepper's (1970:71) eloquent proposition that "to the positivist a hypothesis is a human convention for the purpose of keeping data in order; it has no cognitive value in itself. He is therefore often cynical, or gently indulgent with the wonder and admiration of the common man for scientific predictions". What Pepper suggests is that casually stating hypotheses for the sake of conducting an investigation is not necessarily the correct approach in a positivist paradigm. Good scientific research should therefore start from the basis of established prior theory, which may then provide a logical basis for any stated hypothesis (Creswell, 2002). Since this study was highly empirical in nature, it was appropriate to heed Pepper's *caveat*. By composing specific objectives, while not exclusively relying on pre-stated or contingent hypotheses when insufficient prior theory existed. Thus, overall research goals were ultimately achieved with a combination of validated (when prior theory existed) hypothesised models, statements and specific propositions.

4.2 RESEARCH DESIGN

De Vaus (2003:9) defines the function of a research design as "ensuring that the evidence obtained enables us to answer the initial question as unambiguously as possible". In designing social sciences research, two fundamental approaches are usually followed, namely, positivism or interpretivism. A positivist understands society by using tools typically found in the natural sciences, thereby obtaining rigorous (precise) definitions of phenomena in a contemporary manner by using observable logic to discover causal rules which can be used to predict general patterns of behaviour (Byrne, 1998). This is an important method in the sciences, because testing results' level of replication can challenge or falsify any new theory, model or conclusion. When new and substantive evidence is brought forward, which shows that a prior conclusion may likely be false, the scientific method calls for an adjustment of said theory (Feyerabend, 1985). These requirements were maintained in order to make the present research a highly scientific endeavour. By contrast, the interpretive method involves interpreting society and the behaviour of people in their natural settings, thereby obtaining more fluid definitions of a situation, rather than constructing empirically falsifiable theory (Byrne, 1998). Conclusions drawn in this manner can become highly subjective, and therefore rendered anecdotal. For this reason, the method was not considered for the present research.

This study was dominated by positivism, as it was assumed that the sample group consisted of rational individuals whose behaviour is shaped by their environment. Because the research is a study of a cross-section in the dynamic operational behaviour of airline pilots within a training setting, the approach was eclectic. In other words, various combinations of logical positivism were used in the creation of a hypothesised model of the situation, its operationalization, final analysis and interpretation of the results (for example triangulating empirical data from more than one source).

4.2.1 The research paradigm

Prior to describing the strategy and design of a study, the work must be placed in the context of a particular paradigm that guides the research. According to Creswell (2002), a research paradigm is a philosophical approach to the general nature of the world (ontology) and how we understand it (an epistemology). Similarly, Denzin and Lincoln (2005a; 2005b) explain that an epistemology and ontology are the assumed worldviews adopted by scientists for a particular field of inquiry.

The overall approach taken in this study involved empirical data collection, coupled to a quantitative analytical methodology. Therefore, the specific method used to acquire knowledge from this research was found in *empiricism*, where a structured questionnaire was used as the observation tool of choice. Rationalism was propagated through this scientific method to build any new theory based on the results of the research. Possibly the most effective strategy to gather objective information about the real world, independent of our perceptual knowledge-gathering activities, is the use of the rational scientific method (Feyerabend, 1985). This technique was deemed highly effective for the current research, because human behaviour can be most easily classified, categorised and quantified using statistical methods (Creswell, 2002).

In this study, it was possible to maintain a high level of objectivity because a scientific and methodical quantitative analysis of observed empirical data was undertaken. This entailed a systematic scientific (postmodernist) research design grounded in a theoretical base. As Babbie (2010:10) points out, science "is sometimes characterized as logico-empirical".

For research to be deemed of high quality and useful, it must possess a number of specific properties. The following criteria were followed in terms of Schreiner's (2010) requirements for good scientific research:

- *reference to seminal findings* – by examining the outcomes of previous findings in the field, good research was built on already discovered sound principles;
- *replication in the chain of reasoning* – systematic logic in the reasoning of results improved the quality of the scientific research. When other researchers are able to follow through with the methodology adopted in a particular study, it shows that the findings are plausible, adding to the quality of a study (Muijs, 2004);
- *objective data collection and sampling methodology* – to make inferences regarding the population, mathematical reasoning for selecting the sample were clearly outlined. In addition, the extraction of the necessary data from the sample was conducted systematically and ethically; and
- *concise explanation of phenomena* – to be termed scientific, the final analysis was based on the gathering of observable, empirical, measurable and replicable evidence.

4.2.2 A classification of the overall research design

The following descriptors were selected as best describing the overall design of this study:

- *Empirical research*: In developing a psychological scale to measure perceptions, the study collected and analysed primary data founded in the principles of sociological positivism.
- *Fundamental research*: Findings from the study were not intended to address a specific management dilemma *per se*. The basic aim was to add to the current academic body of knowledge related to the topic. Although fundamental research is born from curiosity, in many instances, it can also provide commercial benefits in the long term (Nelson, 1959).
- *Exploratory and descriptive research*: The design of the study was based on an exploratory premise, because a preliminary literature review revealed that very

little is currently known about perceptions on the advanced automated aircraft training climate (Ausink & Marken, 2005). According to Creswell (2002:16), a researcher conducts a sequential inquiry when (multiple) methods are used for exploratory reasons, followed by quantitative methods "with a large sample so that the researcher can generalize results to a population". The results of the study therefore provided an in-depth description of South African airline pilots' perceptions of the advanced automated aircraft training climate.

- *Cross-sectional research*: Because the survey instrument was only administered to the sample once, it can only provide a once-off or snap-shot view of the theoretical construct. Therefore, an opportunity to examine the structural equivalence of the established scale may exist for future research endeavours.

The overall research design set out a process of constructing and evaluating explanatory statements or theories about the functioning of the real world. Aliseda (2006:6) explains this approach as follows: "[A]n idea leading to a new theory made in science involves a complicated process that goes from the initial conception of an idea throughout its justification and final settlement as a new theory." In determining the steps followed in the scientific theory-building process, one has to differentiate between an adopted methodology that is abductive (where conclusions are based on reasonable estimation), deductive (where conclusions are based on logic) or inductive (where conclusions are based on empirical evidence). This can be viewed as the spectrum of scientific inquiry (that is, abductive-inductive-deductive). In this study, all three methodologies of reasoning were used to draw conclusions.

4.3 REASONING

It goes without saying, that the ability to reason from a set of truths and logical assumptions, is fundamental in order to structure and extract well-formulated conclusions from a scientific research project. Reasoning is the means by which thinking is channelled from one idea to another (Rosenthal, 1994). For instance, Oaksford, Chater and Hahn (2008) discuss the probabilistic approach commonly used in human reasoning, where conditions are set and logical inferences are drawn. In everyday scenarios, a person may, for example, infer that if something is a fish, then it can swim. Therefore, "Nemo can swim" is a logical inference based on a set of conditions and the assumption that the premise is true, that is, "Nemo is a fish" (Oaksford *et al.*, 2008:383). However, it was also borne in mind when conducting the current study that human performance errors are unavoidable at times, and can occur as a result of systematic deviations from logic when the wrong normative standard is used. To take it one step further, logic is expounded through reasoning to derive a valid and particularly substantial means of scientific predictability. For instance, the method was followed in building a predictive logistic model of the phenomena under examination in the present study (see Section 5.10).

The importance of the various reasoning techniques is discussed in the subsections below.

4.3.1 Abductive reasoning

Pierce (1901, cited in Pietarinen, 2006:123) describes the term *abduction* as a logical inference based on estimation (in other words, making a reasonable guess). This leads to the argument that although some truth may be attainable, extremely high levels of certainty may not. Therefore some authors have argued that because high levels of certainty are difficult to attain, more evidence is needed and thus, "abductive foundations are stronger than those based on induction" (Josephson & Josephson, 1996:1). For logic-based abduction, scientists pick out an appropriate explanation or prediction, based on a rational theory representing a domain or set of empirical observations. In other words, by systematically eliminating possible competing explanations from evidential data, the plausibility of the preferred explanation may be supported.

Abductive inference has been slow to develop because logicians concentrate mainly on deductive logic and inductive logic, associated primarily with probability theory. For such reasons, an appropriate statistical level of confidence (in the form of a p-value) is required to substantiate any abductive claims in the observation of human behaviour in social sciences research (Oaksford, *et al.*, 2008). However, the process of abductive reasoning can help guide or steer a proposition or hypothesis by providing the best explanation. For instance, if *D* is a collection of data, and *H* explains *D*, and, in addition, no other hypothesis can explain *D* as well as *H* does, then *H* is probabilistically true (Josephson & Josephson, 1996:14). Although the process is not strongly advocated in the field of psychology and in the behavioural sciences, it was found that the process is nonetheless gaining considerable momentum in attempts to understand complex phenomena involving computer algorithms and knowledge-based systems at an exploratory level (Aliseda, 2006). Josephson and Josephson (1996) have demonstrated the ability to compute explanatory hypotheses without relying on induction, deduction or probability theory. In formulating a plausible predictive model for this research, abductive reasoning was relied upon to select relevant independent demographic variables for testing.

4.3.2 Inductive reasoning

Pietarinen (2006) argues that the inductive process could in fact actually be considered a sub-class of abductive reasoning. According to Feeney and Heit (2007:1-2), inductive reasoning is “probabilistic, uncertain, approximate reasoning”, but is important for the following reasons:

- Inductive reasoning corresponds to everyday reasoning; for instance, we may use induction to predict the probability of what the weather may be like. The empirical evidence collected could be atmospheric pressure, humidity, wind, etcetera, to induce a prediction.
- Inductive reasoning plays a significant role in the behavioural sciences, because it is a multifaceted cognitive activity. Furthermore, it is central to categorisation, similarity judgments, probability judgements, and decision-making.

By means of induction, one can therefore draw generalised probabilistic conclusions from a set of logical empirical observations. In analysing the current data set, statistical methodologies were used to obtain a p-value from empirical observations, and in turn to draw such probabilistic conclusions. The method was also used extensively in content validation and to understand the latent structure of the research construct (see Section 4.13). Moreover, this process of reasoning affords a researcher an opportunity to explore the chance that the possibility exists for the conclusions that are induced to be false, even though all the *a priori* premises may have been true (De Vaus, 2003:85). This quality provided additional robustness to the conclusions that were drawn in the study. Similarly, how well these premises supported the conclusion was based on the degree to which the sample was in general a good representation of the population (Aliseda, 2006).

4.3.3 Deductive reasoning

Deductive reasoning is concerned with "drawing logically valid conclusions that must follow from a set of premises" (Feeney & Heit, 2007: 2). Thus, a deductive argument is deemed the most sound and irrefutable method of gaining knowledge. Deductive arguments are logically valid if the *a priori* premises are true (Aliseda, 2006). Furthermore, this method relies on developing the conceptual or theoretical framework before actual empirical testing. In developing the initial hypothesised research construct it was necessary that the deductive process follow from seminal theory, so as to produce valid results.

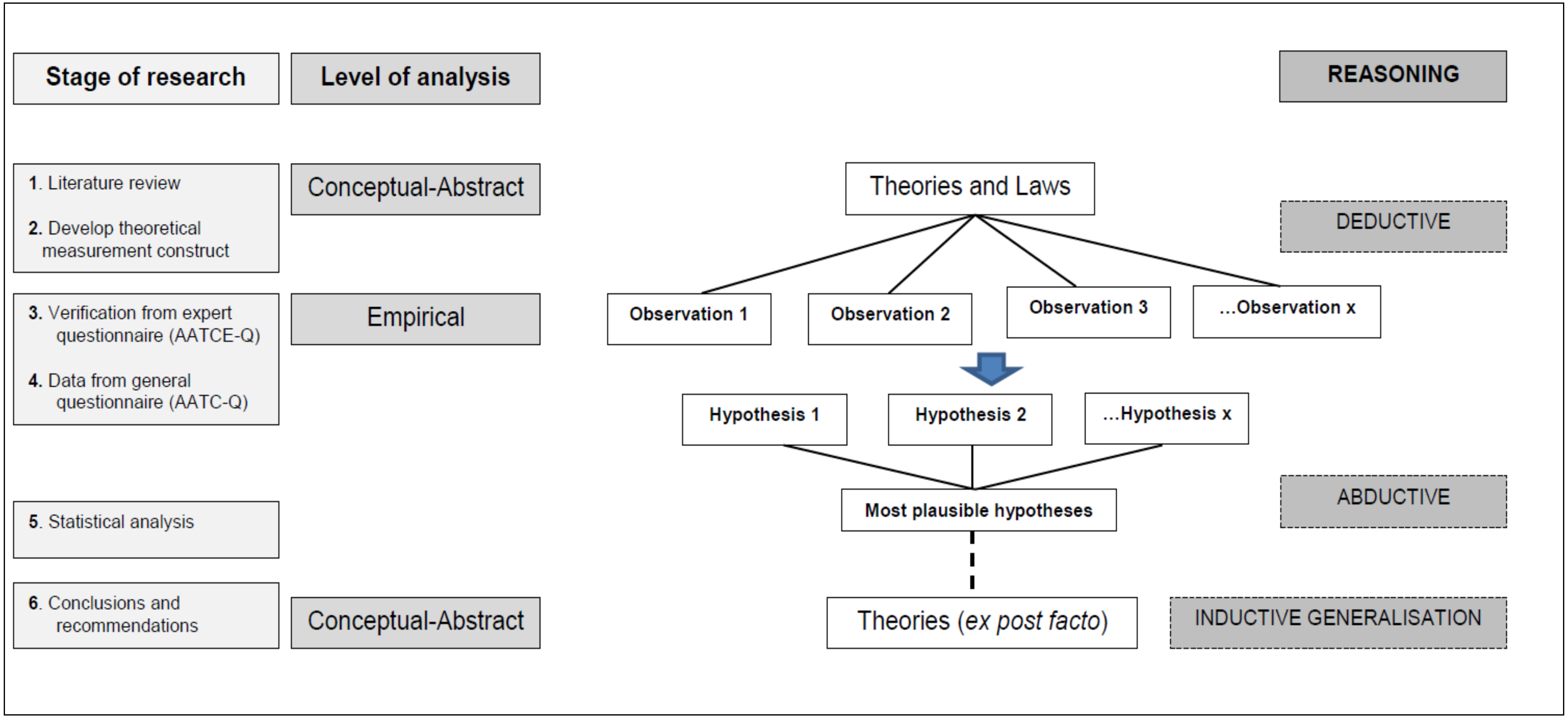
4.3.4 Reasoning followed in the present study

The deductive process has been used successfully in this research to develop an initial conceptualisation of the measurement construct. In the current research, multiple uses of the reasoning spectrum, abductive-inductive-deductive was successfully employed to substantiate conclusions.

Initially, some generalisations were either abduced or deduced from an extensive literature review. The empirical steps of the study involved a combination of both abductive and deductive reasoning, to eventually formulate an *ex post facto* induced theory, and therefore a construction of the final measurement instrument.

Figure 15 contextualises the research design within an eclectic reasoning model. The different stages in the study are contrasted with their level of analysis and the subsequent logical reasoning process.

Figure 15: Integrated model of reasoning used for the study



Source: Adapted from De Vaus (2003), Josephson and Josephson (1996) and Pietarinen (2006)

4.3.5 Ontology

The ontology of a research design has been described in the literature as the philosophical approach taken to the general “nature of existence” (Sullivan, 2009:358). It is a term used to describe categories of entities that may or may not exist in a given domain. For this research, a quantitative ontology was adopted, as it was assumed that the data that was collected could be categorised, classified and counted.

4.3.6 Epistemology

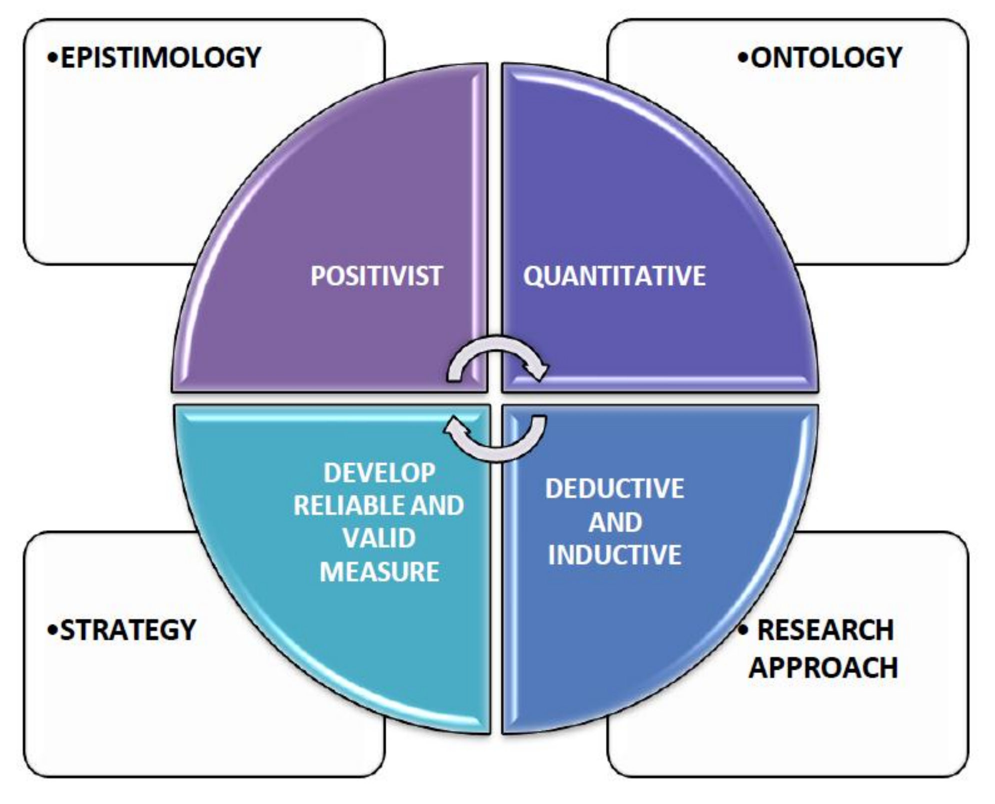
According to Sullivan (2009:180), an epistemology is “the study of knowledge, justification and rationality”. A positivist approach excludes speculation as an appropriate origin for an explanation for phenomena. The knowledge gained from this study is of a positivist nature. It can thus be argued that the syllogistic conclusions derived from this study should be considered an objective truth that has been discovered systematically through observation and measurement.

Since, epistemologists argue that simple belief in a proposition is not substantive enough to be an objective truth, that is, only after observable measurement can one therefore deduce a truth (Nelson, 1959). Therefore, any theory of knowledge can only be considered a truth when a specific belief overlaps observable evidence (proof).

4.3.7 Summary of the research design

In measuring latent socio-psychological constructs (in this case, perceptions of the advanced automated aircraft training climate), developing an appropriate method to operationalize the construct was a core process requirement for the quality of the scale’s construction (Netemeyer, Bearden & Sharma, 2003). A cycle matrix (see Figure 16) can be used to depict the overall design strategy used.

Figure 16: Research design cycle matrix



Source: Author

4.4 THE EMPIRICAL RESEARCH METHOD: A MULTIPLE METHOD APPROACH

A quantitative research approach based on a positivist paradigm and involving the use of a structured questionnaire to gather data from a sample of airline pilots was employed to meet the research objectives. The ultimate aim of the study was to measure the perceptions of the sample after developing an appropriate instrument.

Concepts applied to human perception are not as clear as concepts related to other fields in psychology, making it more difficult to develop a quantitative measurement (Hakala, 2009). Previous researchers probing the perceptions of large samples of technical professionals found that a structured survey method was by far the most effective way of gathering the necessary empirical evidence (Funk & Lyall, 2000;

James *et al*., 1991; Naidoo, 2008; Sherman, 1997). On the basis of the literature review, it appeared that the most appropriate method to follow in constructing a psychological measurement scale of this nature was to apply a multiple empirical method of inquiry.

The study achieved its objectives by relying on a combination of two separate quantitative (positivist) research approaches. This two-step process resulted in the development of the advanced automated aircraft training climate questionnaire. The approach assisted in eliminating human inquiry errors arising from inaccurate observation, as Clark and Watson (1995) also found in their use of such an approach. The initial survey and subsequent quantitative analysis used expert opinion to validate the content of the theoretical construct, in line with Muijs's (2004:2) contention that "quantitative research is essentially about collecting numerical data to explain a particular phenomenon". This step provided a deeper understanding of the construct by explaining the content of the construct.

Some authors have suggested using a combination of both qualitative and quantitative methods in a single study to benefit from the advantages of "triangulation" (Burns & Grove, 2005:226). However, such an approach was not considered feasible for this study, because using both methods is difficult for an investigator – the data that is extracted needs to be interpreted using two very different philosophical paradigms. Nevertheless, it is always theoretically possible to analyse qualitative data in a quantitative manner, for example, by categorising clustered comments from respondents (Leech, 2004). Moreover, in collecting data at a specific level of measurement, a researcher must extract such information from written words, which are language in an extended (con)text, based on observation, interviews or documents. In the current study, as a secondary source of information, words from both the expert and general surveys were loosely analysed to gain clarity and to guide the study objectives (Denzin & Lincoln, 2005a).

The advantages of combining various methods to triangulate data can be harnessed even within a single ontology. Using two surveys to conduct this study achieved this advantage. The technique has been described as adding a three-dimensional quality to the questionnaire approach (Bergman, 2008). This is particularly true of methodological triangulation, which is generally used to analyse complex phenomena (Burns & Grove, 2005). Strict methodological triangulation in scholarly research is usually divided into two main types. The first is *between-methods* triangulation, which is a complex mixture of the qualitative (interpretivist) and quantitative (positivist) paradigms, and is often difficult to accomplish, as mentioned above. The second is a simpler, *within-method* triangulation, which was used in this study. Burns and Grove (2005:227) explain that within-method triangulation consists of a "multidimensional analysis", or the measurement of a phenomenon using two or even three different quantitative instruments. This was accomplished by initially using a subject matter expert probe, followed by the analysis of data extracted from a refined large sample survey instrument (the so-called two-step process adopted).

The final inquiry approach was based on two elements: a *multiple-methods* triangulation (as opposed to a mixed methods triangulation) with a *within-method* triangulation, as also described by Haworth (1996). The final research design therefore consisted of a sequential approach of quantitative methods without blending the different paradigms *per se*, as is the case in many social sciences research projects.

Data were gleaned from multiple sources and interpreted from the perspective of a positivistic ontology. The complications and sources of human inquiry error inherent in the construction of an effective psychological measurement instrument were mitigated by the advantages of methodological triangulation and a multiple-method system.

According to Bergman (2008:91), the advantages of using methodological triangulation are the following:

- *corroboration* – combining methods mutually confirms results, thus providing greater validity;
- *offsetting* – a study is able to take advantage of the strengths found in two separate inquiries by offsetting any of the disadvantages found in either or both;
- *comprehensiveness* – the researcher is able to provide a more thorough account of the field of inquiry by using a multiple step inquiry;

- *instrument development* – clearer and more structured scale items can be devised from a multiple probe of different sources;
- *credibility* – using a multitude of approaches in the inquiry strategy enhances the integrity of findings; and
- *discovery and confirmation* – this implies using diverse views of the phenomenon to generate objectives and employing quantitative methods to confirm hypotheses.

Teddlie and Tashakkori (2008) argue that research evaluation criteria are vastly improved when the intuitive nature of expert judgement is combined with the robustness of a quantitative analysis. The triangulation of diverse positivist methods can therefore significantly strengthen a researcher's inferences.

According to Creswell (2002:16), a multiple-method approach with methodological triangulation offers a study the following research options (applications in this study are briefly indicated):

- *Exploration:*
  By using two quantitative techniques, additional theory regarding the unknown prevailing training climate can be generated from subject matter expert opinion. For instance, in the current study, an expert commented on the issue that new navigational procedures, such as area navigation (RNAV) or precision-based navigation (PBN), could influence perceptions of advanced aircraft training. This was of research interest, because it provides new insight into the complexity of two distinct parts of automation, that is, air traffic control (future air navigation) and aircraft operations.
- *Confirmation:*
  This involves the quantification of separate findings and statistical analysis to test the theory that has been generated. The results of the general survey in the current study could be traced to aspects mentioned in the content validation of the items in the subject matter expert questionnaire. This provided a level of confirmation that could not necessarily be obtained from only the results of a final survey. For instance, in the final survey, it was found that many

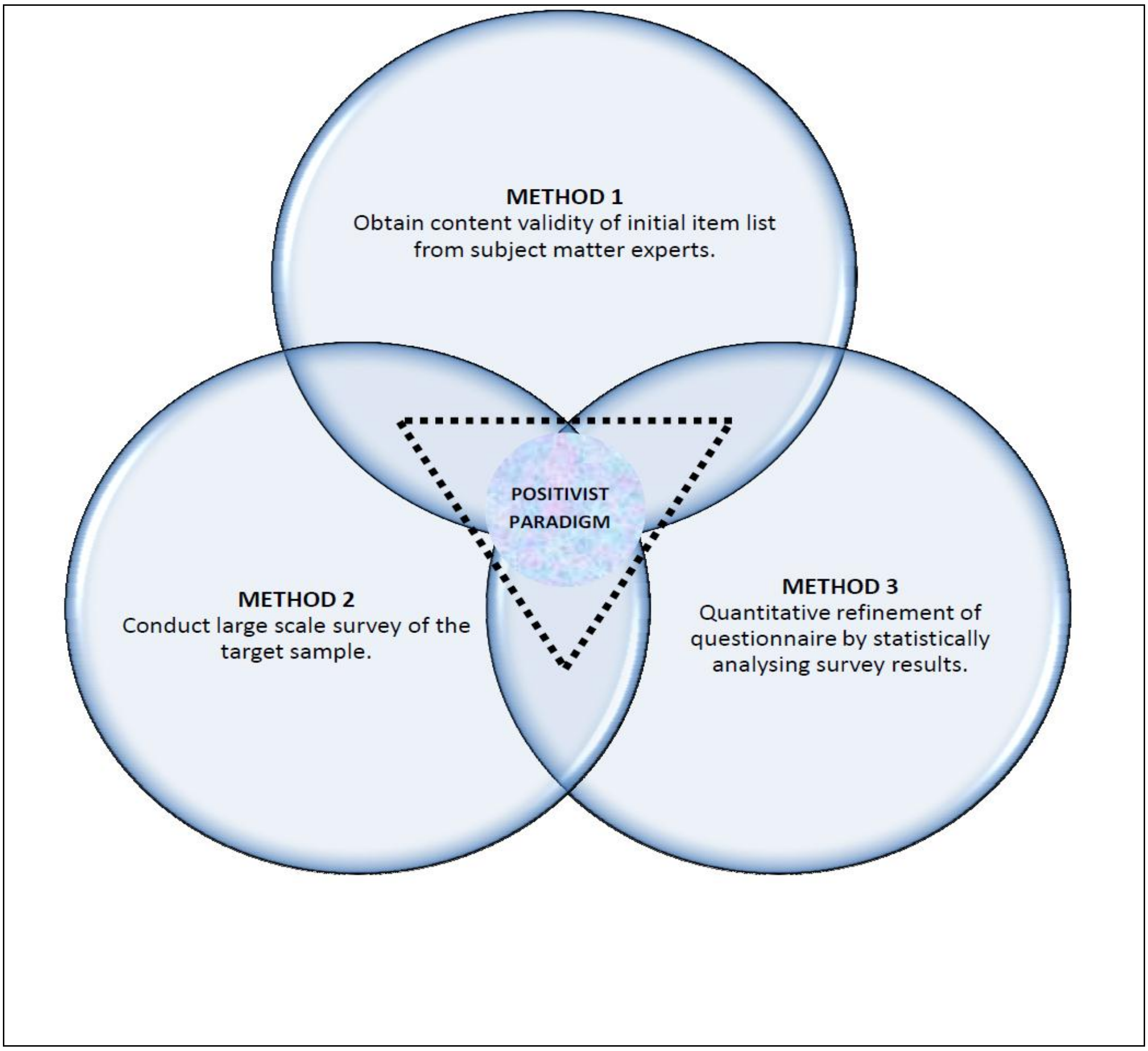
respondents were concerned about the loss of their manual handling skills in advanced aircraft. The experts' comments confirmed this finding, in that some flight instructors were of the opinion that new trainees undergoing transition training to highly advanced aircraft had difficulty in selecting appropriate levels of automation (that is, from fully manual, to fully automated) in adverse or non-normal flight situations.

Figure 17 illustrates the non-linear nature of the research design, which remains in the positivist paradigm. A multiple-method approach exhibits the qualities and benefits of positivist quantification, whilst allowing a researcher to gain from dichotomous, yet similar, methods (Haworth, 1996).

Figure 18 then depicts the *sequential* nature of the multiple-method adopted. Read in combination, Figures 17 and 18 show conflicting event lines (in other words, both circular and sequential), however, far from being paradoxical, the contrasting methods proved highly complementary in contributing to the quality of the final research outcome.

Figure 18 depicts the overall research design, which was divided into the two distinct phases, consisting of four stages overall, which maps the present study. Two *questionnaires* gauged the prevailing perceptions of the advanced automated aircraft training climate construct at each stage. The second quantitative probe was used in developing the final measurement scale.

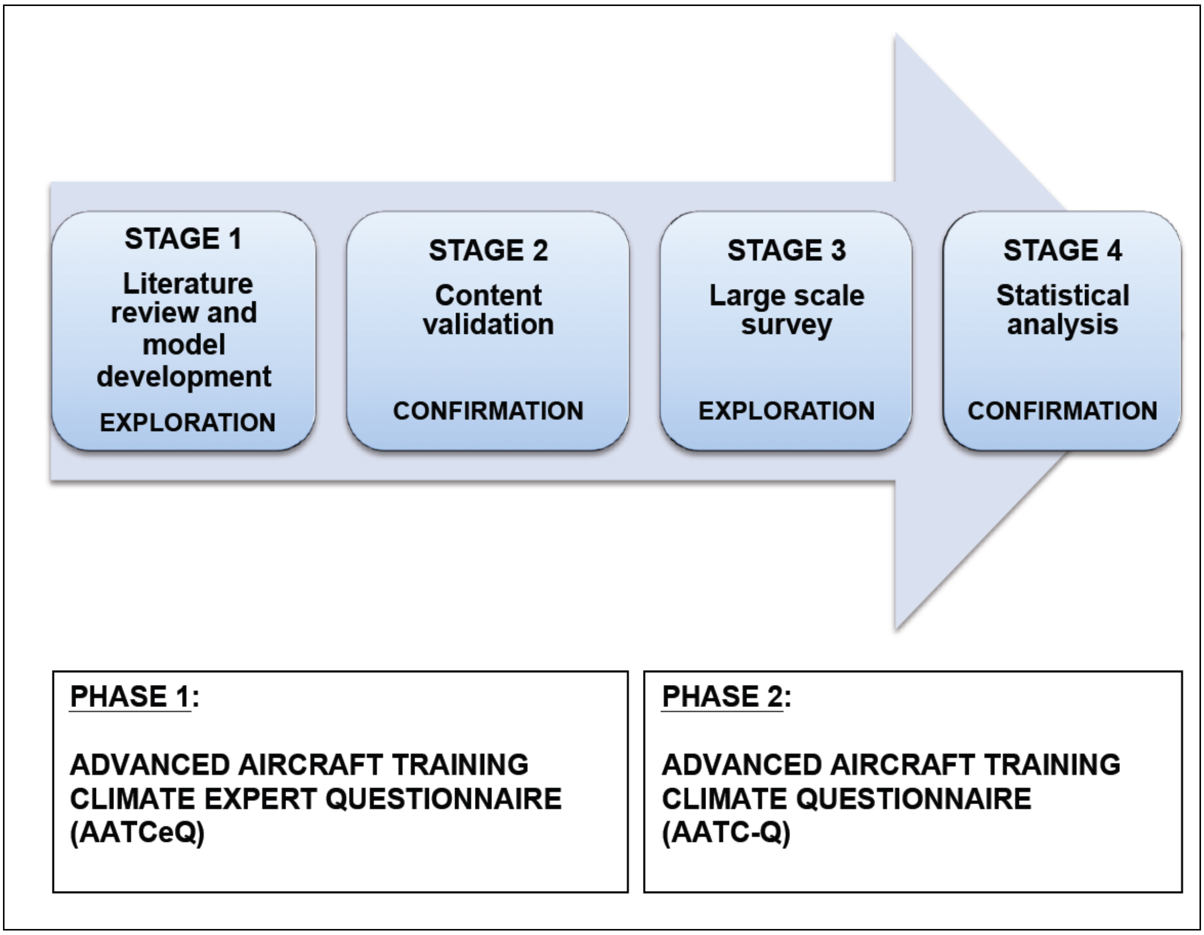
Figure 17: Multiple-method and within-method triangulation



Source: Author

Next, Figure 18 shows a four-stage, two-phase process. The questionnaire in Phase 1 was designed to assess the construct by verifying the relevance, conciseness, clarity, and content validity of the deduced pool of initial measurement items. This was possible, because research questionnaires that are quantitative in nature are traditionally refined on the basis of information derived from previous analyses of opinions (words) gained from earlier investigations (Creswell, 2002).

Figure 18: Overall multiple-method research design



Source: Author

The present case made it possible to further abductively explore textual responses to the expert questionnaire and thereby identify any possible additional variables that could be used in the development of the general survey questionnaire. However, to maintain the authenticity of the content validation process, no new items *per se* were added to the final survey (that is, the content of each retained item was unaltered), resulting in a limited number of very high quality final items. Retained items were however, modified for completeness, clarity and comprehension, as recommended in the textual commentary received from subject matter experts (see Table 23).

4.5 MEASURING INSTRUMENTS

Two primary measuring instruments were constructed to gather the data needed to meet the research objectives. The first instrument was a questionnaire sent to a panel of experts (see Appendix A) to validate the hypothetical construct. The final questionnaire (see Appendix F) was used to survey the perceptions of a sample of the target population, namely, pilots with experience of training for advanced automated aircraft.

As discussed earlier, a preliminary literature review served as the basis for developing a hypothetical model of airline pilots' perceptions of the advanced automated aircraft training climate. The main construct had to be operationalized and measured using empirical evidence, data was gathered by means of a psychometrically valid questionnaire designed to identify latent influential factors. In order to construct a measurement instrument from the initial hypothetical model, 17 specific variables were deduced, at three fundamental levels. All these elements fell within the boundaries of three broad areas of organisational behaviour, which are delineated in seminal works from both classical and contemporary theory.

Items in the subject matter expert questionnaire were constructed for the purposes of testing and validating these critical variables. An item list was generated, based on operationalizing the theoretical construct using abductive and deductive reasoning. The following propositions were formulated to then guide the initial item pool construction:

- Airline pilots' perceptions of the advanced automated aircraft training climate manifest themselves at three levels of organisational behaviour, namely the individual, the group and the organisational levels.
- The theoretical model of the construct can be described in terms of 17 core concepts, namely:
    - learning for technology;
    - motivation to train;
    - personality;
    - training stress;

  - training decision-making;
  - training group dynamics;
  - intergroup training behaviour;
  - training teams;
  - training conflict;
  - power;
  - communication;
  - training culture;
  - knowledge environment;
  - structure;
  - training policy;
  - training standards;
  - training planning.
- The demographic characteristics of the sample differ regarding each of the identified criteria derived from the model, thus indicating various levels of the construct.

The tentative item pool used in the Advanced Aircraft Training Climate Expert Questionnaire (AATCe-Q) consisted of 106 positively worded statements, as Barnette (2000) and Gorsuch (1997) recommend. The validation and analysis of the item pool is discussed in Section 4.13. The final items for the general survey (AATC-Q) were retained or discarded based on the significance of Lawshe's (1975) content validity ratio (CVR).

4.5.1 Survey method

Surveys generally fall into two broad categories: questionnaires or interviews. It was decided that relying exclusively on the questionnaire survey method would prove the most effective way to meet the objectives of the study. Cooper and Schindler (2003:325) suggest that the survey method for collecting data be used when one wants to gain “quantitative information about particular phenomena”. Creswell (2002) suggests that the survey method be used for comparisons and associations, so as to explore whether relationships between phenomena are present. Generally, a survey is conducted on a fairly large scale, as opposed to a laboratory experiment (which is conducted on a much smaller scale). Cobanoglu, Warde and Moreo (2001) point out that, for the purposes of social surveys, questionnaires, interviews and attitude scales can accurately measure participants' perceptions.

The use of a questionnaire to elicit data from a sample may seem intuitive, however, there are a number of disadvantages associated with this method which did in fact prove challenging for the current study. Welman and Kruger (1999) mention some of these disadvantages:

- There is a possibility of a low response rate. This was a real concern, because "[g]etting pilots to participate in surveys is a general problem in the aviation industry all over the world" (Vermeulen, 2011).
- The researcher has a low level of control over the conditions under which the questionnaire is completed. In this case, because both a web survey and hardcopy questionnaire were distributed, there was a real risk that the survey could possibly be completed by inauthentic (or wrongful) recipients.
- Explanation and clarification of concepts is not possible, because space is limited in questionnaires.
- Anonymity complicates the follow-up on questionnaires. Providing a space for respondents to enter an e-mail address if they wished to receive feedback somewhat mitigated this disadvantage in the current study.
- The survey method is generally used for cross-sectional studies, with mainly closed-ended questions. This can be a disadvantage, because exploration of the phenomena under review may be limited.

The rationale for adopting a questionnaire survey approach for this research despite the above disadvantages was based on two very fundamental advantages found in the technique (Welman & Kruger, 1999):

- a lot of information can be collected within a short time span, thereby saving time; and
- data coding is simplified because the survey is structured and standardised.

Two different questionnaire survey methods (hardcopy and electronic) were used to elicit the data necessary for meeting the research objectives. The hardcopy method consisted of a paper-and-pencil survey, whilst the electronic method was based on either e-mail or hosting on the internet (web-based). Both these methods were used to survey the panel of subject matter experts and the final sample frame.

4.5.2 The paper-and-pencil survey

Traditionally, much psychological and management research (unlike research in other scientific fields) makes extensive use of paper-and-pencil surveys to measure abstract theoretical constructs in order to explore underlying organisational phenomena (Schriesheim *et al*., 1993). The advantages of the respondent anonymity that can be achieved using this method have been demonstrated in many studies employing this method (Bradburn, 1983). Participants who opted to use this response method were able to record their answers in a questionnaire booklet at any time and without the potential anxiety of having to answer to an interviewer. Schriesheim *et al*. (1993) warn, however, that the quality of measuring instruments may be reduced when there are such high levels of anonymity.

Despite potential disadvantages, due the nature of the target sample (including the fact that they work shifts), the paper-and-pencil questionnaire proved highly useful in gaining adequate coverage of respondents. The general nature of the work involved in operating a commercial aircraft implies that airline pilots do not occupy a traditional office or always work during conventional times. Hence, access to Internet facilities may be limited. However, in order to improve response rates, both an e-mail questionnaire and an internet-hosted questionnaire were constructed.

4.5.3 Electronic surveying

Apart from the advantages of saving time, convenience and coverage, using computer-based questionnaires also eliminates “out-of-range” responses (Bradburn, 1993:333). Such questionnaires allow only pre-determined valid codes to be entered by the respondents, preventing them from marking inapplicable items. Therefore, the use of this type of questionnaire made the management and analysis of the data much easier for the researcher.

To maintain validity, the paper-and-pencil questionnaire was replicated electronically (see Appendix F). First, an electronic questionnaire was constructed, using Microsoft Word's *form* program, and it was then e-mailed to all eligible participants. Secondly, the questionnaire was re-constructed using an open source online survey application, www.limesurvey.org. The advantages of using a web-based survey are legion. For example, the enhanced import and export functions allow a researcher to use statistical and graphical software far more easily than traditional paper methods would (Nunnally & Bernstein, 1994). For this reason, returns from both the pen-and-paper and e-mail questionnaires in the current study were recaptured onto the web-based survey.

Approximately 64% of completed returns came directly from the web-based survey. In making this research choice, during the construction phase, the drawbacks and advantages of web-based surveys were considered, as set out in Table 13.

Table 13: Contrasting the pros and cons of Internet surveys

Advantages/Benefits	Disadvantages/Drawbacks
The researcher is able to tally results instantaneously, as participants submit responses.	Obtaining the correct sample is not an exact science and can become costly or time-consuming.
The ability to conduct a number of surveys over time is enhanced.	Converting a paper-based survey into an electronic format is time-consuming.
It is easier for respondents to remain anonymous.	It takes both research skill and a fair amount of technical ability to conduct a web-based survey.
The turnaround time from drafting a survey to final execution is shortened.	While an internet survey should be compatible with most browsers, the technology is far from perfect, and can result in increased non-response bias.

Source: Cooper and Schindler (2003:340)

4.6 QUESTIONNAIRE CONSTRUCTION

The final data collection instrument was called the *Advanced Aircraft Training Climate Questionnaire (AATC-Q).* In order to partition the instrument, a demographic section and three core dimensions formed the final design, namely:

- Part A (at an individual level);
- Part B (at a group level); and
- Part C (at an organisation level).

This provided a logical flow of the items and created rapport with respondents.

The study set out to develop a measurement *scale*, as opposed to an *index*. Streiner (2003) points out that the items in an index are an important criterion, but that this is not the case in a scale. This was considered in the design of the general instrument. Items in an index are uncorrelated, whereas in an instrument based on a scale design, in general, items tend to be correlated. This scale attribute also suggested that items should be placed in specific and logical groups. Any correlation between items implies that what one item may miss is usually covered by another item. Because the number of potential items capable of reliably tapping a construct is infinite, the researcher has to choose items appropriately. This ensures that as much of the domain is covered as possible and not just one part of it (Comrey & Lee, 1992). In this case, the researcher was confident that the choice of items selected for the general questionnaire would be a valid measure of the theoretical domain, due to the quantitative technique adopted during the first phase of the scale development (computing the content validity ratio from subject matter expert opinion).

The selection of a correct scale is paramount in shaping the questionnaire and the information collected (DeVellis, 2003). The scales used in survey research usually consist of between two and ten points (or categories), depending on how the data collected is intended to be used (Netemeyer *et al.*, 2003; Stevens, 1946). Debate continues regarding the exact number of points that is best for a measurement scale. Arguments against high granularity suggest that respondents cannot discriminate finely enough to justify more than seven points (Bott & Svyantek, 2004).

An objective in the generation of scale items is to have at least “twice as many items” as the final number needed (Nunnally & Bernstein, 1994:128). The current research generated an initial pool of 106 items, guided by a framework derived from the theory that was reviewed. This number was deemed conservative, because some authors have suggested that around 40 items would be appropriate to measure a construct of this nature (Biggs, 1987; Sherman, 1997). Similarly, Nunnally and Bernstein (1994:130) propose that “at least 30 items” are required for a psychometric measure to have a high level of reliability. Items were revised, and some were discarded as unnecessary items after the Lawshe (1975) analysis (see Section 4.13).

A synthesis of the guidelines (see Table 14) followed for developing a perception scale illustrates the fundamental process used by many authors in the literature. The following generic steps guided the development of a quantitative estimate for the theoretical construct of interest:

- Step 1: Develop a theoretical model of the construct;
- Step 2: Generate appropriate items from the theory;
- Step 3: Operationalize the theoretical construct by developing a scale (for instance, using the results from an expert questionnaire); and
- Step 4: Evaluate the robustness of the scale (appropriate statistics to determine validity and reliability).

The development of the instrument for this research was intended to assess the three key perceptual dimensions of the construct (respondents’ perceptions at the individual, group and organisational levels). Alternatively, the developed measurement’s sub-scales assessed concrete variables, which are related to respondents’ perceptions. According to DeVellis (2003), using sub-scales to divide the number of items in a questionnaire allows a researcher to use fewer respondents for factorisation (in other words, fewer than the 300 required for successful factorisation). This was taken into consideration when the response rate turned out to be lower than expected. According to a rule of thumb provided by Cooper and Schindler (2003), the number of respondents in a sampling frame appropriate for a data reduction method is generally five times the number of items in the sub-scale.

The conception or creation of an initial pool of items is a critical stage in questionnaire construction. Clark and Watson (1995) recommend that researchers err on the over-inclusive side of item generation, so as to derive a broader and more comprehensive item pool which goes beyond the researcher's own theoretical view. The design of items used in the questionnaire in this study is discussed more fully in Section 4.6.2.

To guide the questionnaire construction, Table 14 was used to contrast some important recommendations as discussed in the relevant literature.

Table 14: Contrast of scale development guidelines

DeVellis (2003)	Netemeyer *et al.* (2003)	Pett *et al.* (2003)
1. Determine clearly what must be measured.	1. Clearly define the construct and determine its content domain.	1. Clearly identify the measurement framework.
2. Generate an item pool.	2. Generate measurement items.	2. Identify the empirical indicators of the construct.
3. Determine the format for measurement.	3. Judge measurement items.	3. Design and develop the instrument.
4. Have initial item pool reviewed by experts.	4. Design appropriate study to develop the scale.	4. Pilot-test the instrument.
5. Consider inclusion of expertly validated items.	5. Refine the scale.	5. Determine the number of subjects.
6. Administer items to a development sample.	6. Finalise the scale.	6. Administer the instrument.
7. Evaluate the items.		
8. Optimise scale length.		

4.6.1 Scaling procedure

Fiske (2009:449) comments that it “has been said with justification that the history of science could be written in terms of advances in instrumentation”. Furthermore, according to Netemeyer *et al*. (2003), scaling refers to the measurement of a theoretical construct on a multi-item basis. A latent domain is tapped by using a number of alternative items (scale), providing quantitative estimates of the corresponding construct (DeVellis, 2003). Pett, Lackey and Sullivan (2003) contend that in developing a psychological scale, the researcher is more interested in the construct the items endeavour to measure than in the items themselves. For this reason, in the present study, it was important first to validate the quality of the degree to which the items tapped the construct, prior to developing the actual scale. This was achieved by using the technique advocated by Lawshe (1975), as discussed in Section 4.13.

The most appropriate method for extracting the data needed to measure the construct of interest (perceptions of the advanced automated aircraft training climate) was to use a multi-dimensional questionnaire or instrument containing Likert-type (polytomous) items (Likert, 1958; Pett *et al*., 2003). The items are considered continuous in nature and, in this case, were based on two extreme anchors. The advantage of this technique is that a Likert-type design assumes a latent (continuous) variable with a value that characterises respondents’ attitudes (Likert, 1958). The underlying dependent variable varies quantitatively, as opposed to qualitatively. This is an important quality, which makes the method a popular scale in psychological and behavioural research for measuring opinions, beliefs or attitudes (Cooper & Schindler, 2003; Creswell, 2002; Pett *et al*., 2003). However, Uebersax (2006) found rampant confusion about the use of Likert-type scales and items in many scholarly articles. With this in mind, Uebersax (2006) pointed out that researchers should take cognisance of the following characteristics that have come to define a Likert-type item-based scale:

- the scale itself consists of several items;
- options are arranged horizontally;
- response options are anchored with consecutive integers;

- the response options should, in addition, be anchored with verbal labels representing evenly spaced gradation;
- response options are symmetrical about a neutral point, which implies that the scale should contain an odd number of responses to induce a natural central point; and
- the scale measures levels of agreement or disagreement in respect of a given statement.

Gerbing and Anderson (1988) found that a respondent's behaviour in complying with the internal consistency of Likert's criterion would tend to exhibit a linear and continuous relationship to the score, making it advantageous to use it for statistical analysis. Exploring the latent structure of a construct provided by Likert-type items provides a more robust factor analytic option than other alternatives, such as Thurstone's approach to scaling (Andrich, 1978; DeVellis, 2003).

4.6.2 Item design

In constructing a perception or attitude measurement instrument, "items tend to be very narrow and specific, developed to match a particular situation" (Pett *et al.*, 2003:15). Thus, Kline (2000a, 2000b, cited in Pett *et al.*, 2003) points out that the *quality* of the items tapping the domain of interest is a far more important criterion early in the exploration than psychometric virtues such as validity or reliability. The quality of a scale's inter-item correlations depends, to a large extent, on the number of response options in an item using a Likert-type design (Streiner, 2003).

In deciding on the number of response options that may be appropriate to this study, the researcher followed the steps recommended by Pett *et al.* (2003) and by Gerbing and Anderson (1988):

- Step 1: Decide on what number is appropriate, depending on how well subjects are deemed to be able to discriminate meaningfully between response options relating to statements. Since typical advanced aircraft airline pilots have many years of experience, it was assumed that they have mastered their skill to some degree of expertise. Therefore, potential respondents were assumed to have the ability to discriminate on each item at a far deeper level than the average layman.

- Step 2: Determine whether or not the sample is able to distinguish a construct finely.
- Step 3: Decide how precise the responses should be.

The literature review revealed very little consensus regarding the optimum number of response points to include in an item in a multidimensional Likert-item questionnaire (Streiner, 2003). There are also various advantages and disadvantages to offering respondents an even- or an odd-numbered item scale (Cooper & Schindler, 2003). Creswell (2002) claims that an even-numbered item scale forces subjects to either agree or disagree with the statement, but that this may lead to frustration or even to their discarding the questionnaire altogether. By contrast, an odd-numbered scale may entice some respondents to neglect careful consideration of the statement and continuously give neutral or middle responses (DeVellis, 2003). Nevertheless, several authors, including DeVellis (2003), Field (2009) and Streiner (2003), present convincing arguments in support of the use of an odd number of response options in psychological instrument development. Uebersax (2006) advises researchers to use a neutral point in the design of a Likert-based item, because this method has the advantage of mitigating respondents' frustration levels at not being able to choose a middle stance when they may be unsure of their decision.

The literature review suggested that the majority of perception and attitude measures which used an odd number of item categories in a Likert (1932) design multi-item scale demonstrated very high levels of reliability (Gliem & Gliem, 2003). Furthermore, according to Masters (1974), research findings have shown empirically that for respondents whose opinions do not diverge widely (a relatively homogeneous sample of respondents), the internal consistency of a measure improves as a direct function of the number of categories employed in the item.

With the above argument in mind, a seven-point Likert-type item measuring scale was designed for the present study. For each statement in the scale, the respondents indicated the degree to which they disagreed or agreed with the item. Therefore, high scores would indicate that a respondent held a positive perception of the construct. An example of the seven-point anchored item used in the general survey is depicted in Figure 19.

Figure 19: Seven-point Likert-type item

STRONGLY DISAGREE	MODERATELY DISAGREE	SLIGHTLY DISAGREE	NEITHER AGREE OR DISAGREE	SLIGHTLY AGREE	MODERATELY AGREE	STRONGLY AGREE
1	2	3	4	5	6	7

Source: Adapted from DeVellis (2003) and Likert (1932, 1958)

4.6.3 Rationale for using only positively worded items

Acquiescence bias is another issue of contention when designing a questionnaire. A handful of scholars have demonstrated that a balanced range of items will prevent difficulties encountered by researchers when respondents tend to one extreme of the item scale (Billiet & McClendon, 2000). However, Kristovics (2010) pointed out that these arguments engage in “statistical play”, and are therefore not based on truly scientific reasoning. Kristovics (2010) suggests that researchers should instead maintain a pool of unidirectional statements in scale development.

Welman and Kruger (1999) found that negatively worded items create the error of central tendency. To eliminate this bias, researchers should endeavour to avoid statements that reflect extreme negative positions, for instance “I *never* enjoy simulator training”. Negative items may also tend to frustrate participants or result in their abandoning the questionnaire altogether, especially if the sample comes from a professional population who take pride in the topic under review (Pololi & Price, 2000). Furthermore, such sentiments are corroborated in statistical practicality. For example, an exploratory study by Vermeulen (2009) which used a bi-directional item pool to survey flight instructors’ attitudes towards gender issues required changing the final items to reflect perceptions in a more logical manner – the researcher had to recode negatively worded items so that high scores related to positive perceptions, while the “inverse would be true for low scores” (Vermeulen, 2009:131).

Another danger of using bi-directional items is that such statements may not be true opposites of each other. One can then argue that a truly *balanced* item pool may be very difficult or impossible to construct. Analysing internal consistency, factor structures and other statistics when negatively worded items are used, either together or separately, can be problematic for any researcher (Barnette, 2000; Kristovics, 2010; Vermeulen, 2009). Additionally, in "situations in which respondents can be expected to provide reasoned responses and are willing participants, the need for such a practice would seem to be minimal and may actually be detrimental to the validity and reliability of survey scores" (Barnette, 2000:362).

A primary objective of the study was to determine the underlying structure of the research construct, based on a sample of highly experienced automated aircraft pilots. Participants in the research sampling frame were regarded as professional, as they all hold the necessary licences and certificates as regulated by the Civil Aviation Authority of South Africa, which can be obtained only after acquiring the mandatory levels of training and experience stipulated (CAA, 2011). For this reason, designing positively worded quality items was highly appropriate. The recommendations of Barnette (2000) also played a decisive role in the decision to use only positively worded item statements for the final scale.

4.6.4 Rationale used in the clustering of questionnaire items

Various authors have discussed the advantages of applying item response theory to the structure of scales and the exploration of datasets (Hambleton & Rogers, 1989; Meijer & Baneke, 2004). The origins of item response theory can be traced to the seminal work of Lawley (1943) and Ledyard (1966). Cronbach (1942:299) defines a "response set" as the tendency for a participant to agree or disagree with an item statement, independently from its content. According to Goldstein and Wood (1989:140), "[i]tem response theory (IRT) hinges crucially on the assumption that only a single latent trait underlies performance on an item". This assumption raises the question of whether grouping items into themes or clusters can influence or bias participants' responses. Determining the level of inherent response bias designed into a questionnaire is important, because research into item response theory indicates that the way the items in a survey are constructed can significantly influence the quality of statistical computations.

The order in which to place statements in a questionnaire remains a contentious issue (Simon, Little, Birtwistle & Kendrick, 2003). For instance, Ballinger and Davey (1998) propose a funnelling approach, where questions become progressively narrower in scope. By contrast, Wilson and McClean (1994) suggest grouping statements or questions with a similar topic coverage. The literature also provides examples of the necessity of keeping sensitive statements only in the middle of the questionnaire so as to avoid participant embarrassment early in their participation (Walker, 1996). To achieve the objectives for this research, the questionnaire items were placed in dimensional groupings, as suggested by Wilson and McClean (1994). Items were grouped according to their levels of analysis, that is, either at the micro (individual or person), meso (group) or macro (organisational) level.

The importance of the ordering of the items and its impact on response bias should not be underestimated (Simon, *et al*., 2003). Ordering is important because it presents a contextual effect, which may or may not influence the responses to particular items (Hambleton & Rogers, 1989). Therefore, the design of the current study's structured questionnaire involved clustering items to counter this kind of bias by not *labelling* the underlying or grouping theme. That is, respondents were not explicitly made aware of any particular grouping. Furthermore, only positively worded statements were placed in each latent group to limit the likelihood of bias between the dimensions of the questionnaire.

The main hypothesised construct in this study is systemic, and is therefore comprised of three core dimensions, or sub-constructs (an organisational behaviour approach analyses variables at an individual, group and organisational level). The operationalization of the construct resulted in a questionnaire that extracted both demographic and perceptual data (opinions, beliefs, attitudes). The aim of the Phase 2 questionnaire development (the AATC-Q) was to produce statements that represented each of the three dimensions and 17 conceptual themes identified from the literature study. A structured questionnaire with clustered or grouped items was deemed the most appropriate method for capturing such perceptual data.

The structuring of the survey provided respondents with alternatives to each question in a Likert-type item. The current items of the questionnaire were captured in three latent themes (the individual, the group, and the organisation). Clustering items according to such themes in scale development is adequate because, according to Bejar (1983), dimensionality is situation-specific. This means that dimensionality is not purely a property of the items itself, but rather a response to the items under a specific set of conditions. This approach resulted in an accurate assessment of the latent structure (see Section 5.2.4). Therefore, it can be said that a response set provides better data when it remains in its natural thematic setting, as opposed to being randomised (Wilson & McClean, 1994).

Alternatively, a review of the literature revealed that there are as many reports of no or trivial order effects as there are of significant or important order effects – “[a]t present, therefore, the frequency, size, and nature of question-order effects in standard surveys of the general population are matters of considerable uncertainty” (Schuman & Presser, 1996:24). The decision to maintain underlying themes from the order of items for this research was based on the original intent of organisational behaviour analysis, which implies measurement at three distinct levels. Furthermore, maintaining specific themes or dimensions within a questionnaire is based on the premise that the groupings themselves admit items that are *only peripherally* related to the underlying unitary theme. Fundamental to item theory is the notion that psychological constructs are “latent” (Meijer & Baneke, 2004:354). The perceptions of these constructs can only be obtained from the manifest responses from participants to a set of items. According to Meijer and Baneke (2004), the structure of a research questionnaire assumes the existence of a latent trait on which persons and items have an opinion or take a position. In the current study, this implied the need to group items in line with the assumption of the existence of latent themes (traits) prior to a factorial data exploration. Randomisation of items may have dissolved the assumed structure. Item clustering then provided a more accurate description of what the variables were actually doing and, more specifically, acknowledged the nature of organisational behaviour theory as *substantive*. In addition, it will be observed that the results of the factor analysis (see Section 5.2) revealed a latent underlying structure of the items which themselves correlated across latent themes. Factor analysis was the *statistical* method of choice, which determined that clusters of items were actually related to one another. Furthermore, according to Goldstein and Wood (1989:164), "unidimensionality in the presence of multidimensionality will produce a composite dimension".

The item grouping choices made for the purposes of the current research can be summarised according to the position of Schuman and Presser (1996), who argue that grouping similar questions together presents a smoother organisation of the questionnaire and appears sensible or coherent to respondents. The negative effects of any ordering sequence are far too inconclusive to warrant a randomised set of questionnaire items. In addition, many scholars are fairly confident that major findings in their research were not due to response order effects as such in any case (Schuman & Presser, 1996).

4.7 STRUCTURE AND LAYOUT OF THE QUESTIONNAIRE USED IN THE STUDY

The overall structure and layout of a questionnaire has been known to influence the responses participants are willing to give, as well as the overall response rate (Cooper & Schindler, 2003). An introductory letter (see Appendix B) was therefore attached to each paper-based questionnaire, and a similar introductory letter preceded the web-based survey (Appendix F). The main body of the questionnaire was highly structured. This entailed that alternatives were provided to the respondents, who had the simplified task of marking only the appropriate answers. The purpose of the questionnaire was to elicit data from the sample with regard to their demographic particulars and their perceptions of, or attitude towards, advanced automated aircraft training. The questions found in most of the questionnaire were closed-ended, and took the form of Likert-type items. According to Babbie (2010:256), closed-ended questions can be "easily" processed and provide for better uniformity of responses, as opposed to the alternative, which is open-ended questions.

Table 15 depicts the layout of the final questionnaire (AATC-Q).

Table 15: Questionnaire structure

Section	Topic of section	Number of questions
A	Demographic information	22
B	Perceptions of the advanced automated aircraft training climate	42
C	Participants' comments and feedback	2
Total number of questions		66

Section A consisted of questions related to the demographics of each participant. Specific questions referred to the person's age, gender, educational qualifications, levels of experience as a pilot in terms of years and hours, type of aircraft operated, perceived level of computer literacy, and whether the person had enjoyed his or her most recent flight simulator and route training experience.

The airline pilot's experiences, opinions and perceptions of their training were then gauged in Section B. Each perception statement was presented as a seven-point Likert-type item. The items were also clustered according to the level of measurement at a micro, meso and macro level. To limit any response bias associated with clustered items (see Section 4.6.4), the various categories of analysis (micro, meso, macro analysis) were not indicated to the respondents in the questionnaire itself.

Section C provided an area for the participant to interact with the researcher if the participant wished to do so. Comments by respondents were recorded here. Textual data is an important source of information collaboration and can be used to verify or clarify ambiguous findings. Participants were also given an opportunity to provide their e-mail addresses for future correspondence on the study results and to communicate any interesting recommendations. This option was intended to allow the possibility of providing feedback and close the knowledge loop (closure).

4.8 LEVELS OF MEASUREMENT

Many researchers have had difficulty in deciding whether data extracted from items in a Likert-type design are “continuous, categorical or rank ordered” (Stevens, 1946:677). Clason and Dormody (2001) argue that it is highly probable that the summated items from a Likert-type designed questionnaire are ordinal or interval, and thus approximate a continuous scale. In addition, it is generally assumed that an ordinal or interval Likert-type item is continuous, because, according to Nunnally and Bernstein (1994), behavioural research scales measuring perceptions assume an approximately equal interval scale with considerable assurance. Since the study was intended to measure the perceptions of the automated aircraft training climate construct and its associated variables, the study was designed to measure airline pilots’ attitudes by means of an interval scale (Likert-type, continuous data).

Likert (1932) originally constructed five-point items in a summated scale to assess survey participants’ attitudes. However, Likert (1932) admitted that the number of intervals in the item might be open to manipulation, and subsequently no fixed number of intervals was recommended in the original Likert-type item. The confusion in the literature regarding Likert-type scales and Likert-type items still persists. Clason and Dormody (2001) point out that Likert’s seminal work was not intended to develop a summated scale in the first place, although the questionnaire items appeared as a scale of some sort. With this in mind, in the current study, individual seven-point Likert-type statements were adopted in which a rating from 1 to 7 implies varying levels of disagreement or agreement with the statement (that is; strongly disagree, moderately disagree, slightly disagree, neither agree or disagree, slightly agree, moderately agree, strongly agree).

A level of measurement stems from the granularity of the items (number of intervals). High granularity is based on the assumption that respondents in a sample can discriminate fairly accurately due to their enhanced levels of experience in the given field (as was the case among the respondents in the current study).

Cooper and Schindler (2003:223), Morgan, Leech, Gloeckner and Barrett (2007:42), describe the characteristics of four different types and levels of measurement, in terms of “ratio”, “interval”, “ordinal” and “nominal”. However, variables’ levels of measurement were originally contemplated by Stevens (1946:678) to clarify and determine the nature of data. Such clarification improved computational quality by guiding the selection of appropriate types of statistics to be used to explore the data further. The details pertaining to the levels of measurement according to Stevens (1946) are:

- *Nominal* (categorical scale) measurement is used for the empirical determination of equality, as in gender (male or female). Permissible statistics for this measurement level are the number of cases, the chi-square, McNemar, phi or Cramer’s V, and discriminant analysis.
- *Ordinal* (rank ordered scale) measurement is used for the empirical determination of greater or lesser value, as in the perceived quality of training received (very good, good, average, poor). Permissible statistics for this measurement level are rankings, mean rank, median and mode. The Mann Whitney-U, Kruskal-Wallis, Spearman, rank order correlation or Kendall Tau are preferred measurement tests.
- *Interval* (continuous scale) measurements are used for the empirical determination of scores that are ordered from low to high in categories that are evenly spaced. For example, a summated Likert-type designed scale of which the items measure on a “strongly agree” to a “strongly disagree” continuous seven-point scale would be considered an interval level measure. Permissible statistics for this measurement level are mean, standard deviation, factor analysis, Student’s t-test, one-way analysis of variance (ANOVA), Pearson’s correlation, regression analysis, multiple regression analysis, factorial ANOVA and multivariate analysis of variance (MANOVA).
- *Ratio* (continuous scale) measurements are used for the empirical determination of the equality of ratios, as in most physical measurements, for example, age in years or hours of experience in advanced aircraft. These measures have equal intervals between the levels or scores and a true zero level. The permissible statistics used for each measure are cumulative, in other words, all operations discussed above can be used for calculations involving ratio type data.

4.9 RESEARCH POPULATION AND SAMPLING STRATEGY

The “universe” of elements in which a researcher happens to be interested is commonly referred to as a “population” (Butcher, 1966:3). Many scholars stress the importance of defining the population correctly, because doing so determines the level of the statistical accuracy of the final sample. For this reason, Cooper and Schindler (2003:181) propose that the “ultimate test of a sample design is how well it represents the characteristics of the population it purports to represent. In measurement terms, the sample must be valid”. The validity of a sample is highly dependent on its accuracy (absence of bias) and precision (degree of error).

The target population for this study consisted of individual persons, in particular, qualified South African airline pilots who have some level of experience with advanced commercial aircraft.

4.9.1 Determining the sample size

Determining a sample size that makes it possible to extract sufficient data for statistical analysis can be a difficult exercise for researchers. In determining the most suitable sample size, three criteria are usually specified (Kalton, 1999):

- the level of precision required (for the social sciences, an acceptable level of error is 3%);
- the level of confidence or risk accepted (in social sciences research, an alpha level of 0.05 at the *a priori* level is acceptable where the *ex post facto* effect size is evaluated); and
- the degree of variability in the attributes being measured (designed using Likert-type items, provided a level of continuous data).

Bott and Svyantek (2004) have suggested two fundamental reasons for ensuring an accurate sample size for conducting scientific research:

- a minimum number of cases is required to analyse sub-group *relationships* adequately. For factor analysis (discussed in Section 4.17.5) around 200 cases are required if several items are used to define each construct); and
- in order to draw associational and comparative conclusions, the sample must, as far as possible, *represent* the population under scrutiny.

Two separate sampling procedures were conducted during the study. The first step called for experts to provide statistical validation of the questionnaire items of the training climate dimensions and their descriptive elements, which were initially identified theoretically. For a classical statistical analysis of expert judgements, Lawshe (1975) strongly suggests a minimum of 15 panellists for quantitative validation.

The second part of the research relied substantially on the results of an exploratory factor analysis. It was noted that DeVellis (2003) claims that a large number of unspoilt returns (around 300) are required for a factor analysis to be reliable (in other words, to uncover dimensional clusters). There are, however, other opinions on this topic, so it was explored further to determine the most appropriate path to follow in order to obtain a workable sampling frame for the study.

An analysis of the literature presented conflicting and varying propositions on determining the most appropriate number of elements to provide a good sample (see Table 16). For instance, Stoker (1981) suggests that the sample size should be proportional to the number of elements contained in the population size (N), whereas Welman and Kruger (1999) argue that, irrespective of the size of the population, it is not necessary to use a sample larger than 500 units for the analysis. This suggestion is in line with findings reported by Gravetter and Wallnau (2008), who have demonstrated that the standard error in a sample size is reduced exponentially, and not in a linear fashion. Therefore, the standard distance between a sample mean and the population mean tends to be reduced with larger samples, although it will never drop to zero even for extremely large samples.

Table 16 synthesises some important authors' sample size requirements for the development of a valid and reliable psychometric measurement instrument. It is, furthermore, important that the sampling method adopted be reported accurately, so that readers can draw their own conclusions (Bartlett, Kotrlik & Higgins, 2001). It is clear from the comparison of methodologies that a sample of 200 to 300 observations is adequate to provide a stable factor solution for the instrument. Comrey and Lee (1992) suggest that 200 elements in a sample is a fair to adequate number for obtaining relatively stable solutions.

Table 16: Contrasting notions of what constitutes a good sample size

Source	Recommendation
Stoker (1981)	Proportional to /N (for example, when N=1000, minimum sample size=141)
Arrindell and Van der Ender (1985)	20 times the number of factors
Comrey and Lee (1992)	100:poor; 300:good; 1 000:excellent
Nunnally and Bernstein (1994)	10 observations per variable
Welman and Kruger (1999)	Not necessary to have more than 500 observations
DeVellis (2003)	At least 300 observations required to conduct factor analysis
Netemeyer *et al*. (2003)	5 to 10 observations per parameter estimated
Pett *et al*. (2003)	10 to 15 subjects per item
Tabachnick and Fidell (2007)	A minimum of 300 cases
Saunders, Lewis and Thornhill (2007)	278 cases in 1 000 will provide a 5% margin of error

Alternatively, using Cochran's (1954) sample size formula for scales based on seven-point Likert-type items, Bartlett *et al*. (2001) calculated that for a finite population of around 1 400 elements (which was the target population for the current study), the required sample size was only 118, and corrected to 111 (when the sample size exceeds 5% of N).

Oversampling may be necessary when return rates are expected to be low, as is typical with survey research of this nature. Therefore Bartlett *et al.* (2001) determined that for a population of 1 400, it may be necessary to send out a minimum of 171 questionnaires if the anticipated return rate is estimated at 65%. With this in mind, in addition to the electronic means used for data collection, a total of 700 hardcopy surveys were distributed to ensure a good response rate. This suggestion also requires a sample of no less than 5% of the population for acceptable accuracy and precision in social sciences research.

Stoker (1981) and, more recently, Streiner (2003) recommend that researchers bear in mind the three boundary constraints when considering the size of the sampling frame (that is, level of precision, confidence interval and degree of variability). The main concern for researchers determining the ideal sampling size for an exploratory factor analysis stems from sampling error (Osborne & Costello, 2004). However, when item communalities (the amount of variance explained by common factors) are relatively high (0.6 and above, as was the case in the present study's data set), sampling error is somewhat reduced, and an exploratory factor analysis produces a fairly stable factor solution using smaller frames of between 200 and 300 elements (MacCallum, Widaman, Zhang & Hong, 1999). Moreover, in their analysis, Osborne and Costello (2004) found that neither the number of variables nor the size of *N* had any significant unique effect when all other variables were kept constant. The levels of variable communalities were high in the present study because the scale development began with content validation from subject matter experts.

4.9.2 Sampling frame based on the response rate

The actual cohort of the final sample in the current study was related to the response rate. The response or return rate is usually expressed as a percentage (the ratio of the number of questionnaires sent out divided by the number of usable questionnaires returned). A multitude of factors may influence the response rate. Haworth (1996) suggests that around half the final response will be obtained without the need to send participants a reminder. Another third of the responses can be obtained from a first reminder. This technique was used to elicit additional returns from participants in the current study too. The technique resulted in a response rate of approximately 33%. In social sciences research similar to that in the current study, the average response rate was found to vary around 30% (Osborne & Costello, 2004). Therefore the response rate for the present research was satisfactorily typical.

After reviewing the results of a rigorous versus a standardised survey methodology, a response rate of 33% was not completely disappointing. To answer the question of "[w]hat differences arise in point estimates subject to different response rates", Keeter *et al*. (2000) compared two surveys. In their study, they completed two surveys: a rigorous survey conducted over five days, which obtained a 60.6% response rate, and a standard one, which obtained a 36% response rate (in other words, nearly half the rate obtained in the five-day survey). Perhaps surprisingly, Keeter *et al*.'s (2000) study found that, despite the differences between the two survey responses or return rates, both achieved very similar statistical results. Their survey with a lower response rate was only minimally less accurate than its more rigorous counterpart. However, it was nonetheless borne in mind that variances increase when samples are smaller than the target number of minimum returns (Bartlett *et al*., 2001).

In summary, Haworth (1996) provided some very important methods to obtain good response rates and reduce non-response bias. These include

- concentrating on the design of the questionnaire (careful layout);
- maintaining a logical ordering of questions;
- clear phrasing of statements, combined with an attractive presentation; and
- endeavouring to keep participants interested in the topic to elicit greater participation.

4.9.3 Sampling procedure

According to Kalton (1999), in order to conduct replicable scientific research, it is necessary to clearly state and implement definitive statistical reasoning when selecting only some elements from a population. Therefore, in order to draw conclusions, make inferences or devise theories about a population, one must have a mathematically sound basis. Researchers usually use two fundamental sampling scheme categories, first, random or probabilistic, and, second, non-random or non-probabilistic (Cooper & Schindler, 2003). A probabilistic method requires that each element of the population frame have an equal chance (a non-zero probability) of being selected for inclusion in the final sample. This method requires an accurate list of the elements in the population and can prove expensive. A non-probability method was used for this study, based on the fact that a list of the entire population was unobtainable. A guideline on the actual numbers of eligible pilots was, however, obtained from both the Civil Aviation Authority and the Airline Pilots' Association of South Africa. These numbers were used to determine an appropriate size for the sampling frame.

The judgement, quota, snowballing or convenience sampling methodologies are examples of the most common non-probabilistic methods used in similar research (Creswell, 2002). A purposive judgemental method was used in the current study, based on the guidelines, using the pilot numbers from the Civil Aviation Authority and Airline Pilots' Association. In addition, after interviewing and then using the judgement of experts who are particularly knowledgeable about the field and phenomena under study, due consideration was given to the systematic inclusion and exclusion of certain elements from the population, as recommended by Babbie (2010). In order to extract a representative sample for the South African situation, the population was stratified according to the various major airline companies based in the country. Cooper and Schindler (2003:193) describe such a stratification method as partitioning the population into mutually exclusive "sub-populations or strata".

The primary unit of analysis was the perceptions of *airline pilots*; hence, the target population consisted of only those South African airline pilots who held a current licence to operate advanced automated aircraft at the time of the survey. According to the figures provided by both the Civil Aviation Authority and the Pilots' Association of South Africa, the population was estimated at approximately 1 400 pilots.

A non-probability method was used to gain access to a convenient sample. Questionnaires were purposefully distributed to the stratified groups of individuals in accordance with Haworth's (1996:47) suggestion. The probability of selecting a particular entity from the sub-population for the sample frame using this method was unknown in terms of the criteria proposed by Bott and Svyantek (2004). In other words, systematic randomisation (where each entity has a known non-zero chance of selection) was not obtained. According to Kalton (1999), non-representativeness is a distinct disadvantage when using such sampling techniques. Obtaining the required sample size by targeting elements in a stratum of interest offsets some of the disadvantages found in the non-probability sampling technique and provided a level of control and precision, as described by DeVellis (2003). In this case, elements in the population of interest (airline pilots) could be regarded as highly homogeneous by nature, with very little significant variation in opinion, as was the case in a prior similar study (Naidoo, 2008). This premise also reduced sampling error in the final sample frame, because “how large a sample should be is a function of the variation in the population parameters under study” (Cooper & Schindler, 2003:190).

The non-probability, convenience and purposive *stratified sampling* technique (Cooper & Schindler, 2003) used in this study entailed dividing the population into several strata or groups. Stratification is the process of partitioning members of the population into relatively homogeneous subgroups before sampling (Kalton, 1999). In this case, the homogeneous strata were based on the specific airline company to which each element belonged. Saunders *et al*. (2007) suggest that convenience sampling be used when there is very little variation in the population, as was the case in this target population.

The population itself was deemed to contain little variation, because it is common knowledge that all airline pilots employed at major carriers are selected only after a battery of tests, and after complying with the certification requirements stipulated by the South African Civil Aviation Authority (SACAA). These tests serve as a filtering mechanism for each organisation to ensure that only those candidates who fit the corporate culture of the particular airline are selected. Hence, it was reasonable to assume that the *source* of the data was limited to a fairly homogeneous cross-section of qualified airline pilots flying advanced automated aircraft in various South African airlines.

In order to target specific strata, a number of airline organisations were also approached for assistance in maximising the response rate. These organisations are regarded as the largest airlines in South Africa and fit into the Airline Pilots' Association of South Africa (ALPA-SA) portfolio, namely (also see Table 17):

- South African Airways (SAA);
- British Airways Comair (BA Comair);
- South African Express Airways (SAX);
- Mango Airlines (Mango);
- South African Airlink (Airlink); and
- 1Time Airlines (1Time).

4.9.4 Stratification in terms of airline pilot unionisation

To explore other stratification options, pilot unionisation was considered, because, amongst the airline pilot group in South Africa, unions play a major role in organisational perception. Also, to maintain some level of anonymity for the organisations under study, it was decided to partition the six participating airlines into groupings according to whether the pilots were unionised or not. The major carriers in South Africa can easily be separated into organisations, which have large numbers of unionised pilots (membership of more than 60% of the pilots employed at the organisation) on the one hand, and those which do not on the other hand. The population was partitioned in this manner to allow for easier categorical comparisons. Airline pilots tend to gravitate towards those organisations that boast larger numbers of unionised members due to perceived improved working, training and safety standards (Walsh, 1994). Such perceptions may also have a significant influence on opinions of the training and overall organisational climate (Olney, 1996).

Airline pilot unions are considered separately from traditional industrial unions. They are generally considered professional bodies and are regarded more in terms of an association (ALPA-SA, 2011). In the case of some smaller airlines, airline pilots may be represented by large industrial unions, such as Solidarity. Be that as it may, the current numbers of job applicants are higher at airline companies with ALPA-SA membership, and labour turnover at the airlines without such representation is higher. In South Africa, the legacy airlines report the largest number of unionised members (ALPA-SA, 2011). At the two oldest and largest airline companies in South Africa (SAA and BA Comair), at least 99% of the pilots are unionised. Higher salaries, pension and provident funds, coupled with a significantly better safety record and a non-punitive organisational culture appear to be the primary attraction (Walsh, 1994). Olney (1996) postulates that structured unification of professional employees improves training standards and subsequently organisational climates, because many professional associations regard themselves as an integral part of efficient enterprises.

According to ALPA-SA (2011), currently, half of South African airline organisations are unionised and half are not. SAA, SA Express and BA Comair account for the bulk of the unionisation, while the smaller carriers – South African Air Link, Mango Airlines and 1Time Airlines – are non-unionised companies.

4.10 BASIC DEMOGRAPHIC INFORMATION ON THE FINAL SAMPLE

Aaker, Kumar and Day (1995) propose that the representation of the population within the sampling frame has more significance on post analytical results than the actual response rate in itself. In addition, Cooper and Schindler (2003) also point out that relying on sheer magnitude from numbers, would not guarantee a representative sample. A primary disadvantage from using a convenience sampling method, is that population representation within the sampling frame is compromised. However, by targeting specific sections of the population of interest, selection bias was to a certain degree, mitigated.

The decision to utilise an Internet based survey method resulted in a level of unavoidable under coverage of the target population, leaving certain demographics underrepresented. It was hoped that by using a hybrid data collection method (paper-based and web-based surveying), the adverse effect of the Internet for surveying, would be reduced.

Table 17 shows that, in general, the population was well represented. The majority of the participants in the sample frame (48.7%) are positioned within the organisation employing the largest number of airline pilots in South Africa (approximately 800 pilots). The concept of representation is especially important when a stratified sampling method has been adopted, as in the case of the current study.

In addition, the desired categories were well represented, apart from gender (see Table 17). The distinct inequity in the distribution of male and female pilots is nonetheless an accurate reflection of the current status of the aviation industry, as there are very few female airline pilots. Previously, South African legislation prevented potential female candidates from pursuing a career in aviation, but change, albeit slow, is now occurring at many airlines. However, because the current study was not focused on gender issues or gender phenomena as such, the skewed distribution within the gender category was not regarded as an aggravation in terms of the analyses of results.

Table 17: Respondent sample frame (N=229)

VARIABLE	FREQUENCY	PROPORTION	MEAN (S.D)
ORGANISATION			
1 (SAA)	112	48.7%	
2 (BA Comair)	23	10.0%	
3 (SAX)	14	6.1%	
4 (Airlink)	34	14.8%	
5 (Mango)	11	4.8%	
6 (1Time)	10	4.3%	
7 (Other)	25	10.9%	
SIZE OF AIRLINE COMPANY			
Large (1+2)	135	58.5%	
Medium (3+4)	48	21.4%	
Small (5+6+7)	46	20.1%	
MAIN AIRCRAFT MANUFACTURER			
Boeing	57	24.9%	
Airbus	95	41.5%	
Other	77	33.6%	

Table 17: Continued

VARIABLE	FREQUENCY	PROPORTION	MEAN (S.D)
GENDER			
Male	212	92.6%	
Female	17	7.4%	
AGE (years)			41.28 years (11.359)
Below 30	38	16.6%	
30 – 40	81	35.4%	
41 – 51	56	24.5%	
52 – 63	51	22.3%	
Above 63	3	1.3%	
EDUCATION LEVEL			
No tertiary education	131	57.2%	
Tertiary education	98	42.8%	
INSTRUCTOR RATED			
No	102	44.5%	
Yes	127	55.5%	
FLYING EXPERIENCE (hours)			9 753.3 hours (6116.719)
Below 2000	7	3.0%	
2 001 – 5 000	58	25.3%	
5 001 – 7 000	30	13.1%	
7 001 – 10 000	39	17.0%	
10 001 – 15 000	57	24.9%	
Above 15 000	38	16.6%	
COMPANY STATUS			
Captain	120	52.4%	
Co-pilot	109	47.6%	
COMPUTER LITERACY			
Poor	5	2.2%	
Average	87	38.0%	
Above average	92	40.2%	
Excellent	45	19.6%	
INITIAL TRAINING			
Military	81	35.4%	
Airline cadet	18	7.9%	
Self-sponsored (part-time)	64	27.9%	
Self-sponsored (full-time)	66	28.8%	

Table 17 clearly shows that in terms of the general flight experience levels of the group, the sample was fairly well distributed, with the majority of respondents above the 5 000 hour mark (Mean=9753.29; SD=6116.719). However, the dispersion of the participants in terms of flight experience was large – the majority of the sample had between 3 000 and 16 000 flight hours. The high standard deviation of this descriptor is a testament to the heterogeneity of pilots found in the South African airline industry. Most of the pilots with the national carrier regard their present organisation as the final step in their career progression and some will retire after spending almost 40 years there (ALPA-SA, 2011). This is a further indication of high levels of industry experience. This was expected, as the target population were all qualified airline pilots operating advanced aircraft. Airline companies operating such aircraft tend to hire very experienced pilots.

The experience of the group can also be reflected in the mean age of 41 years (SD=11.359). The distribution of the participants' ages ranged from the mid-20s to the late 60s, indicating that, in terms of generational analysis, the airline pilot group is a fairly disparate one. This provided a good area for further statistical analysis, which was then undertaken as reported in Chapter 5 where the age category was sub-divided or combined as required, for an in-depth exploration of the relevant phenomena.

In general, most of the respondents (59.8%) perceived their levels of computer literacy as better than average. It may be hypothesised that by virtue of the fact that participants operate relatively superior machinery, their presumed technological acumen becomes pervasive. Secondly, one major South African carrier provides company laptop computers to its pilots, therefore possibly facilitating improved perceptions of computer abilities and skill within the target group.

The airline organisations in South Africa were further categorised in terms of size. The size of an organisation is generally determined from the sheer number of employees, the market reach or market share it enjoys and the extent of its operations (Desler, 2002; Drucker, 1946). According to Pitfield, Caves and Quddus (2010), a large major carrier is described as operating a fleet of aircraft with the company brand and identity in terms of the ICAO (International Civil Aviation Organisation) or IATA (International

Air Transport Association) code. The major airline also has a unique call sign associated with it. For instance, South African Airways has the call sign "Springbok", whilst British Airways has adopted the call sign, "Speedbird" (IATA, 2012). However, in order to define the various airlines in South Africa in terms of being either large or small, it was considered whether the organisation operates at least one fleet of more than 10 aircraft, which is capable of carrying more than 99 passengers upon its national operating certificate (Child, 1973). As a middling category however, it was necessary that the well-known regional carriers be positioned in the medium size airline group (SA Express and SA Airlink).

Most airline pilots employed at the largest organisation (in this case, South African Airways) operated Airbus-manufactured advanced aircraft (41.5%), which was reflected in the skewed proportions regarding aircraft type and manufacturer category. Appropriate non-parametric methodologies were subsequently employed (see Chapter 5) in the data analysis to understand and further explore the phenomena associated with the aircraft type sub-groupings. Employing more robust statistical methods (non-parametric procedures) mitigated the impact of any adverse effect emanating from the fact that only 24.9% of the participants indicated that they operated Boeing-manufactured advanced aircraft.

Instructors (non-rated or rated), level of education (tertiary or no tertiary) and company status (captain or co-pilot) were relatively well balanced, providing for good comparative examinations later in the thesis.

4.11 DATA COLLECTION PROCEDURES

Cooper and Schindler (2003:87) define data as "the facts presented to the researcher from the study's environment". There are many methods to extract raw data from the field. Such methods include, but are not limited to, questionnaires, standardised tests, observational forms, laboratory notes, and instrument calibration logs (Cooper & Schindler, 2003). Alternatively, collection methods for large-scale surveys include electronic mailing (e-mail), internet-based e-survey submissions (for improved response rates), together with traditional paper-based questionnaires (Cobanoglu *et al.*, 2001).

Because empirical research requires data to be collected, in this case, first from a group of subject matter experts, and thereafter from a number of respondents in the target population, it was decided that a structured self-administered questionnaire would be used in both cases. The description, design and administration of the subject matter expert questionnaire are discussed later in Section 4.13. The advantages of self-administered questionnaires, as described by Cooper and Schindler (2003), include the benefits of expanded geographic coverage, minimal staff requirements, and the use of complex instruments, allowing respondents time to think about questions. The greatest disadvantage, however, was a low response rate (apathy).

Apart from the conventional survey distribution methods currently used in the airline industry, such as box dropping (personal letter boxes), the assessment instrument (AATC-Q) was administered to the sample population via the distribution channels used at ALPA-SA, namely its web page and e-mail contact list. Both the Association's executive committee and the different airline management groups graciously offered their assistance to maximise the response rate. Correspondence regarding the goals and intentions of the research project was communicated directly through email, telephone and one-on-one interaction with both airline management and pilots' association executives, so as to gain the necessary support and endorsement of the present study. In order to ensure an adequate response rate and a greater number of unspoilt returns, the instrument was also hosted on the World Wide Web as a dedicated e-survey that replicated the hardcopy questionnaire.

A cover letter explaining the purpose of the survey (see Appendix B), together with a note of the endorsement from both ALPA-SA and the company's management, accompanied each questionnaire in an attempt to entice participation and therefore improve the response rate. For data collection purposes, both the expert group and the target population were nonetheless readily accessible to the research team.

Subsequent to the development of a draft of the aforementioned large sample survey instrument (the questionnaire), the validation of a pool of items (constructed after an in-depth literature review) was analysed by a purposive group of subject matter experts. An expert in the area of modern advanced automated aircraft training is defined, for the purposes of this study, as an academic experienced in the field of aviation management, or a highly experienced flight instructor (with an advanced licence rating). It is generally accepted in the aviation industry that there is a positive correlation between total flying time and mastery of skill (Sherman, 1997; Telfer & Moore, 1997). To contrast the construct validation, it was necessary that a proportion of the expert panel be current academics (advanced educational credentials) in the field of interest. Therefore each subject matter expert was either an academic in the field, or held many thousands of hours flying instructional experience on advanced automated aircraft. The next section reports on this item validation process.

4.12 CONTENT VALIDATION

The quality of a perception measurement instrument rests on the level of validity in its content (Cooper & Schindler, 2003). The first step in the research plan required the use of expert opinion in refining the derived questionnaire items and thereby obtaining a valid content that could operationalize the construct of interest. It is a challenge to ascertain the content validity of a measurement scale based on the opinions of experts in the field (Landis & Koch, 1977) – Hardesty and Bearden (2004:98) suggest that there is “a lack of consistency and guidance regarding how to use the expertise of judges to determine whether an item should be retained for further analysis in the scale development process”.

After obtaining sufficient data from the judges, there were two areas of potential inaccuracy, which may have affected the quality of the measurement scale. The first of these inaccuracies stemmed from potential *inter-observer bias*, which consists of differences between the marginal distributions of the response variable associated with each of the observers (Altman, 1991; Fleiss, Levin & Paik, 2003; Karlsson, 2008). Cochran’s Q-test, a test in the analysis of variances, was subsequently used to test the hypothesis that inter-observer bias was absent.

The second inaccuracy was *observer disagreement*, which reflects the fact that observers may classify individual items in the same category of the measurement scale. Karlsson (2008) suggests that the computation of the Kappa test statistic coefficient can be used to determine the level of inaccuracy associated with the data set. However, an alternative method based on the value of a ratio calculated according to Lawshe's (1975) formula (discussed in this section) was used to determine the level of agreement between the judges' categorisation of items in the current study.

To determine whether the measure's items actually capture a proper sample of the theoretical content domain, opinions from experts and/or inter-rater agreement were sought, in line with Karlsson's (2008) suggestion. In order to gain a representative sample of the content domain of the unobserved construct of interest, judgements regarding whether possible items may actually represent the intended construct were then validated. According to Hardesty and Bearden (2004), there is some confusion between face validity and content validity, which are terms that are also, to some degree, used interchangeably in the literature. Some authors suggest that expert opinion primarily evaluates *face validity.*

According to Fleiss *et al.* (2003), the *content* validity of a measure can then be validated indirectly through a statistically significant inter-rater agreement calculation. Hence, for the purposes of meeting the research objectives, the first phase of the study evaluated the statistical significance of inter-rater agreement as an indication of the content validity of each item's relationship with the construct and sub-constructs, this was calculated using Lawshe's (1975) method and Cochran's Q statistic.

In an attempt to gain a more in-depth understanding of how the research process acquired content validity, an analogy was constructed. The universal domain of the construct under study is represented by the contents (universe of acceptable items) entering a funnel (which represents the inter-rater or expert judgement filtering procedure). In order to obtain a proper representation of the main construct of interest, items are hypothesised to belong to one of three sub-constructs (consisting of specific item clusters). This was based in accordance to Hardesty and Bearden's (2004) premise that a measurement scale (measuring the main or super-construct) would not have the required content validity if its items accounted for the variability in only one exclusive sub-construct. Hence, if items appeared to fall into the opening of the analogical funnel, it would have face-validity. In other words, according to the expert judgement method, these items were actually measuring the main construct at some level. With this analogy it is easy to imagine how a different researcher measuring the same construct of interest, may obtain different indicators (items) to the ones discovered in this specific study. It is possible that different items can measure (tap) the same construct, because an infinite number of indicators manifest in one variable space.

The purpose of expert validation in this study was therefore to ensure that the items in the initial pool reflected the desired main hypothesised construct of interest. Eventually, after Lawshe's method of item analysis, and a statistical assessment of inter-rater bias, the final item pool consisted of fragments from the universal domain of available items (see Figure 20). The aim of the filtering phase in this scale development was to ask subject matter experts to judge whether,

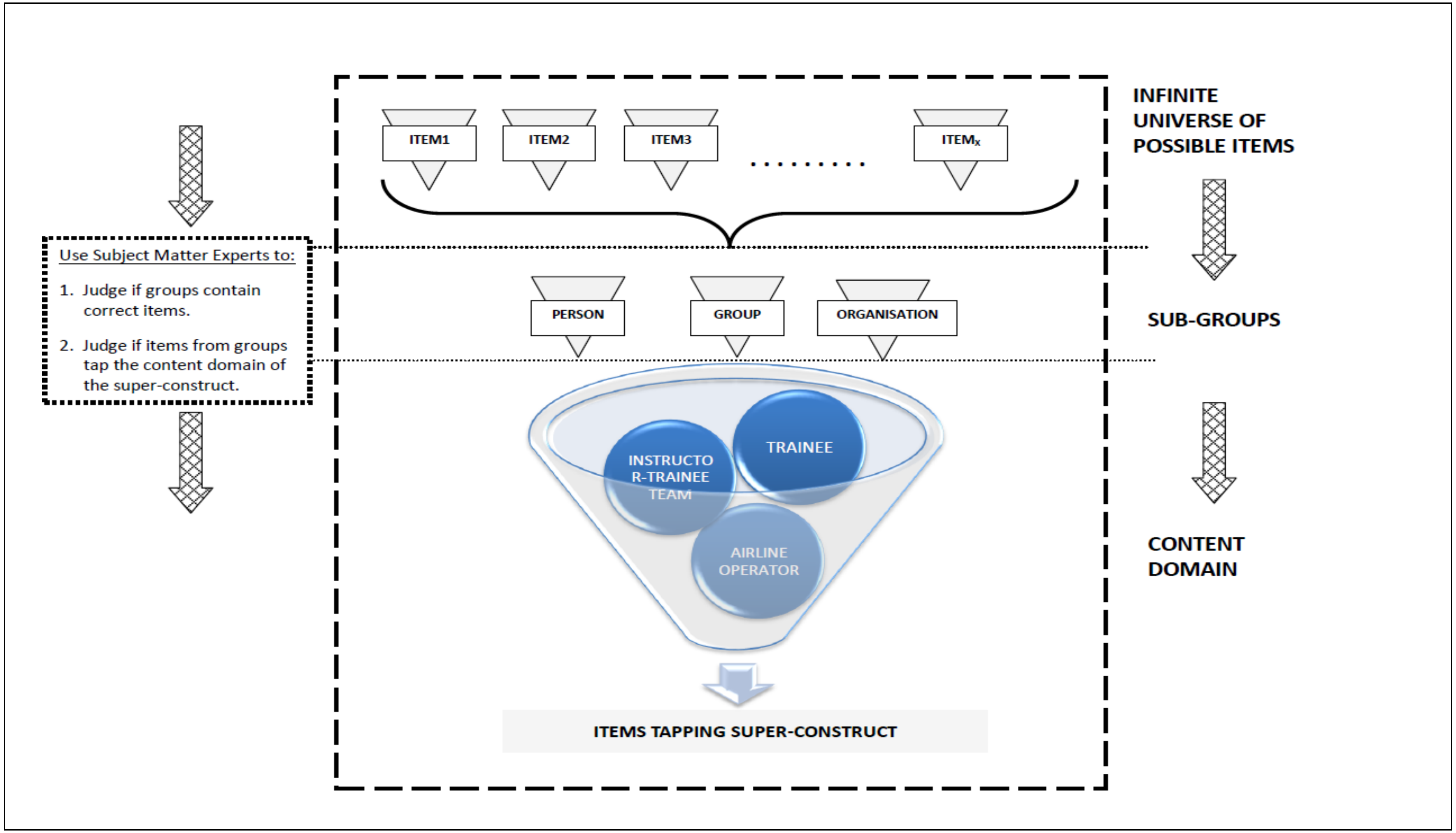
- groups contained correct items; and
- items from groups tapped the content domain of the super-construct.

Analysing expert judgment is a validity process undertaken before data collection, therefore "the development of a new measurement instrument is [generally] a two-stage process" (Karlsson, 2008:110). Altman (1991) warns that there are many problems in causally determining levels of agreement between judges during initial scale development. Inter-rater agreement is a process rife with systematic error. It was found that many researchers conducting scale development use associational statistics incorrectly in an attempt to obtain content validation (Landis & Koch, 1977), therefore it was necessary to also check the level inter-rater bias post-validation.

Fleiss *et al*. (2003) found that the percentage of agreement between judges or correlations determined using a Pearson's coefficient could be highly misleading. These and other reasons prompted the pursuit of a more robust content validation method. To develop a valid and reliable scale, it was decided that the technique proposed by Lawshe (1975) was the optimum solution in the initial stages of scale development.

Figure 20 was developed to propose an analogy that illustrates the process followed to validate the content of the hypothesised construct. Content validation was deemed an important early step in the study, as it created the foundation for subsequent data collection, analyses and final discussion of phenomena.

Figure 20: Content validation analogy



Source: Author

The level of agreement between subject matter experts was based on Lawshe's (1975) method of content validity because the method is regarded as *mathematically* sound. Judges were asked to determine how "essential" an item cluster is to a specific sub-construct representing the content domain. Independent views were elicited from the experts by asking each expert to respond to the following question in terms of the measurement of the hypothesised construct: "Is the knowledge measured by this item cluster: essential, useful but not essential, not necessary?"

Lawshe (1975:567) developed the following formula for the computation of the minimum content validity for different panel sizes based on a one-tailed test at a significance level of $\alpha = 0.05$:

Content Validity Ratio (CVR) = $(n_e - N/2)/(N/2)$,

where:

n_e = number of experts indicating "essential"; and

N = total number of expert panellists.

Lawshe (1975:566-567) suggests that a minimum CVR value of "0.49 is required from 12 to 15 subject matter experts" to ensure that agreement is unlikely to have been due to chance. Alternatively, Fleiss *et al*. (2003) suggest that the statistical value of a Kappa coefficient would also confirm significance (this method was not pursued in the current study). For this study, 36 subject experts were approached to participate, and 17 usable sets of responses to the questionnaires were returned (a response rate of 47%, which was deemed fair, and therefore adequate for the analysis to continue [Streiner, 2003]).

A CVR value of 0.46 is required to obtain the necessary validity when using a panel of 17 experts (Lawshe, 1975). A more conservative cut-off point of 0.49 was however, subsequently used (see Section 4.14).

A non-exhaustive list of 106 items was generated from the literature review to hypothesise the operationalization of a model of the construct. Of the 106 items, 64 were deemed not essential or necessary for having some degree of content validity.

Thus, 39.62% of the original item list was retained after analysis of the opinions from the panel of subject experts. The next section describes these results in more detail.

4.13 RESULTS OF LAWSHE'S TECHNIQUE

A final cohort of 17 highly experienced airline flight instructors and university academics participated in the expert validation process. An instrument in the form of a survey questionnaire was developed to extract data from the sample of experts (see Appendix A). The instrument contained five main sections as follows:

- *Section 1:*
  This part of the instrument contained an introductory letter and information regarding respondent consent. It introduced the research to the expert and provided the contact details of the researchers.
- *Section 2:*
  This part of the instrument contained information about the background literature review on the topic of interest. More importantly, this section of the survey showed the expert respondent what the hypothesised model of the construct consisted of (as discussed in Chapter 3).
- *Section 3:*
  This part of the survey asked for the respondent's demographic information.
- *Section 4:*
  The important data collection statements were contained in this section of the expert survey. This part of the expert instrument was further divided into three dimensions. The first dimension (27 statements) solicited information about the organisational level (the airline) of the construct. The second dimension (27 statements) asked experts about their opinions regarding statements related to the group level of analysis (the instructor-trainee team). Finally, the third dimension (52 statements) in this part of the instrument solicited information about the individual level of analysis on the construct (the trainee), from the expert respondent.
- Section 5:
  The final part of the expert survey was qualitative in nature. Here, experts were

asked about their opinions regarding the clarity and comprehensiveness of the items.

The subject matter expert questionnaires were distributed electronically and in hardcopy format. A follow-up request was made to the experts after two weeks. Due to the length and depth of the subject expert questionnaire, it was difficult to convince participants to complete the request timeously. Of the 36 questionnaires distributed, 17 were returned, giving a response rate of 47%. This response rate is regarded as average for studies of this nature (Streiner, 2003).

Table 18 (demographic data) and Figure 21 (distributions) show that the mean age of the panel was 54.23 years (SD=7.64). The participants displayed a high degree of industry experience, with a mean of 30.65 years (SD=10.82). The mean instructional experience of the airline pilots was 3 780.64 hours (SD=2023.97), indicating a very high level of expertise in the subject.

A minimum of 15 panellists were required to attain a CVR of 0.49 in order to accept an item as essential in tapping the construct of interest. The distributions depicted in Figure 21 furthermore show clearly that the data are skewed. Skewed distributions in this context thereby confirm that the experts come from the tail of a normal curve. This was expected, as subject matter experts cannot be regarded as being in the same category as the average large survey respondent.

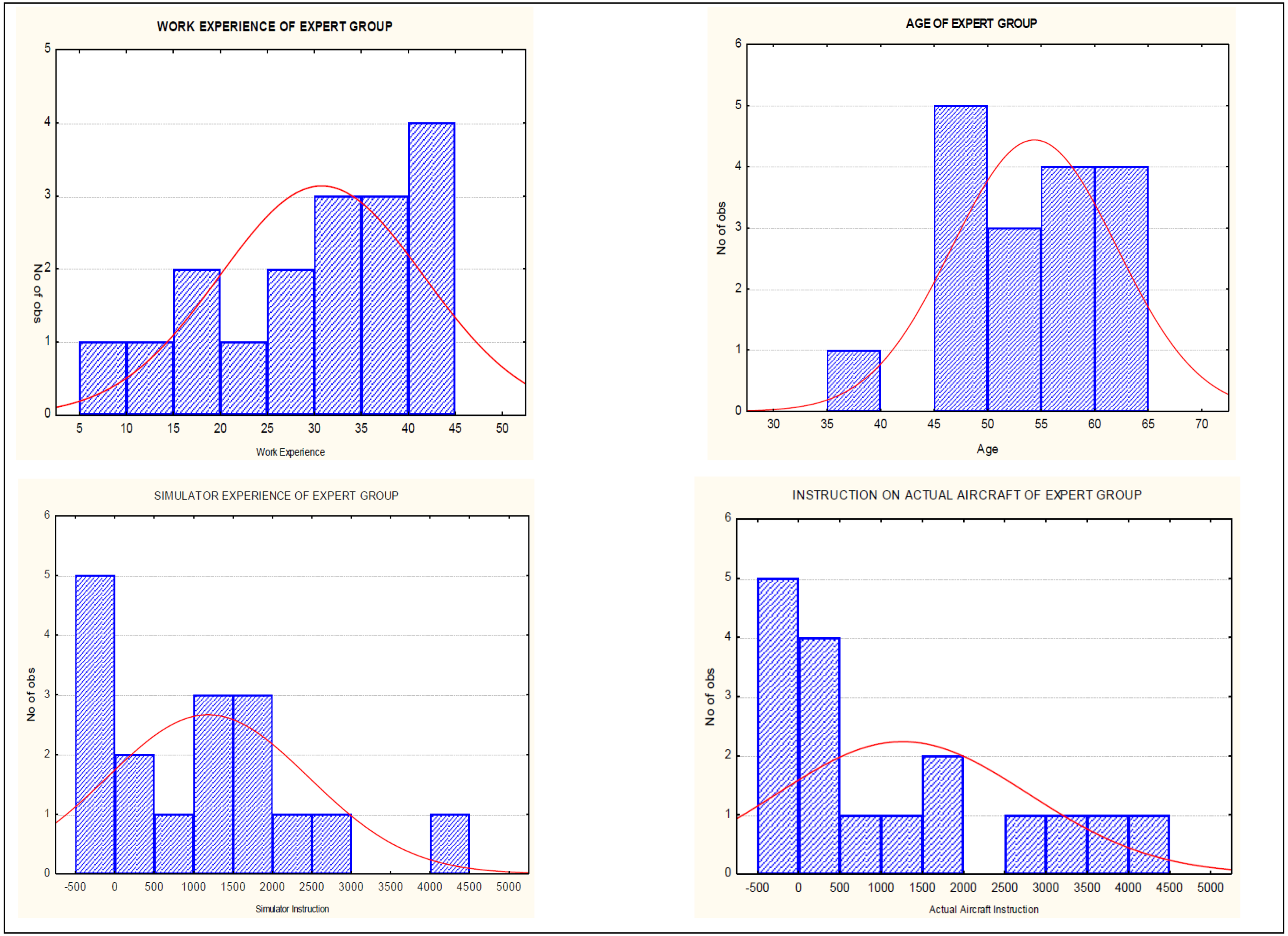
Table 18: Demographic data of the subject matter experts (N=17)

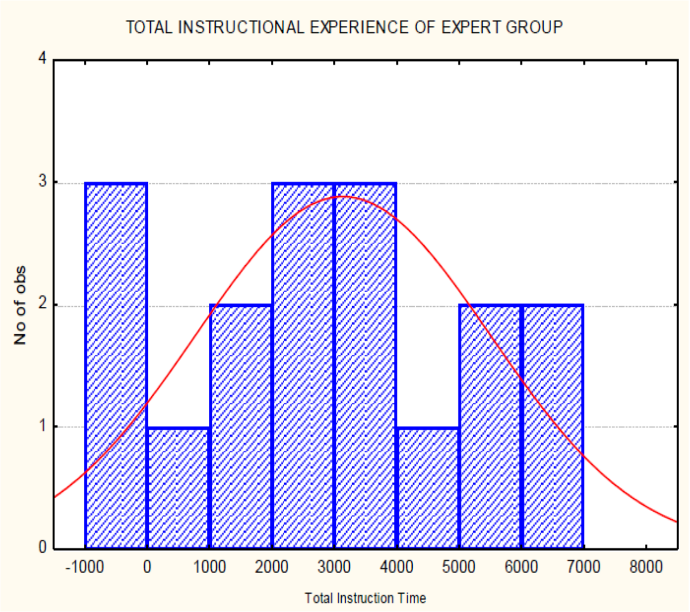
DEMOGRAPHIC VARIABLE	COUNT	PERCENTAGE
AGE (years)		
31-40	1	5.88
41-50	5	29.41
51-60	7	41.18
61+	4	23.53
INDUSTRY EXPERIENCE (years)		
10-14	1	5.88
15-20	3	17.64
21-25	1	5.88
26-30	2	11.76
31-35	4	23.53
36-40	2	11.76
41+	4	23.53
TITLE		
Airline Pilot (training instructor)	12	70.59
Academic	3	17.65
Airline Pilot and Academic	1	5.88
None (indicated)	1	5.88
HIGHEST EDUCATION ATTAINED		
Secondary School	5	29.41
Diploma	3	17.65
Bachelor Degree	2	11.76
Honours Degree	1	5.88
Masters Degree	2	11.76
Doctoral Degree	4	23.54

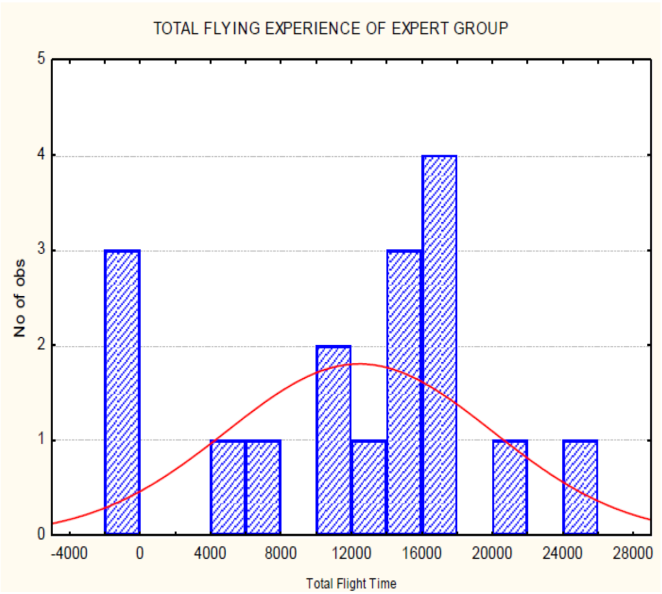
Table 18: Continued

DEMOGRAPHIC VARIABLE	COUNT	PERCENTAGE
FLIGHT SIMULATOR INSTRUCTION (hours)		
0-500	7	41.17
501-1000	1	5.88
1001-1500	3	17.65
1501-2000	3	17.65
2001+	3	17.65
ACTUAL AIRCRAFT INSTRUCTION (hours)		
0-500	9	52.94
501-1000	1	5.88
1001-1500	1	5.88
1501-2000	2	11.76
2001+	4	23.54
TOTAL FLIGHT TIME (hours)		
<5000	3	17.64
5001-10000	2	29.41
10001-15000	4	23.53
15001-20000	4	23.53
20001+	4	23.53

Figure 21: Distribution of subject matter expert demographic variables




TOTAL INSTRUCTIONAL EXPERIENCE OF EXPERT GROUP
No of obs
0
1
2
3
4
-1000
0
1000
2000
3000
4000
5000
6000
7000
8000
Total Instruction Time


TOTAL FLYING EXPERIENCE OF EXPERT GROUP
No of obs
0
1
2
3
4
5
-4000
0
4000
8000
12000
16000
20000
24000
28000
Total Flight Time

Tables 19 to 21 show the results of Lawshe's (1975) technique to assess content validity.

Table 19: Lawshe test results for Domain A

ITEM	ELEMENT	Endorsement of statement		CVR	RETAIN (Y/N) (Reject if CVR < 0.49)
		Essential	Not essential or not necessary		
A1	Pilot training at my airline is in line with company goals.	17	0	1.000	Y
A2	My company's training produces world-class pilots.	15	2	0.764	Y
A3	I have noticed a steady improvement with regard to pilot training at this company.	11	6	0.294	N
A4	I know what my company's training goals are.	15	2	0.764	Y
A5	My company has talented people managing airline pilots' training.	15	2	0.764	Y
A6	Pilot training at this company is professional.	15	2	0.764	Y
A7	Management follows the regulator rules appropriately.	15	2	0.764	Y
A8	Pilot training on this aircraft is well organised at this company.	17	0	1.000	Y
A9	Pilots who are engaged in simulator training are professionally attired.	3	14	-0.647	N
A10	I understand what the company expects of me when I am in training.	16	1	0.882	Y
A11	It is easy to share my training experiences with colleagues at this company.	7	10	-0.176	N
A12	Training at my airline produces safe pilots.	16	1	0.882	Y
A13	There is a well-established chain of authority for pilot training on this aircraft.	12	5	0.411	N
A14	This airline gives its pilots an appropriate amount of preparation work before training.	13	4	0.529	Y
A15	The paperwork involved in training for this aircraft is appropriate.	11	6	0.294	N
A16	It is easy for me to appeal for assistance if I encounter a training problem at this airline.	16	1	0.882	Y
A17	There is sufficient training guidance from the company.	16	1	0.882	Y

Table 19: Continued

ITEM	ELEMENT	Endorsement of statement		CVR	RETAIN (Y/N) (Reject if CVR < 0.49)
		Essential	Not essential or not necessary		
A18	**The standard operating procedures (SOPs) for learning to fly this aircraft are adequate.**	17	0	1.000	Y
A19	**The company provided me with sufficient time to prepare for training on this aircraft.**	17	0	1.000	Y
A20	**The simulators my company uses to train its pilots are in good condition.**	14	3	0.647	Y
A21	**I feel motivated by my airline to train for this aircraft.**	8	9	-0.058	N
A22	**The training department at my company is flexible.**	6	11	-0.294	N
A23	**The airline is very supportive of its pilots' learning requirements for this aircraft.**	16	1	0.882	Y
A24	**My company's culture supports training for new technology aircraft.**	16	1	0.882	Y
A25	**There is sufficient feedback about my training on this aircraft.**	17	0	1.000	Y
A26	**Pilot training at my airline follows civil aviation requirements.**	16	1	0.882	Y
A27	**My company uses only current training material.**	15	2	0.764	Y
	AVERAGE NUMBER OF ENDORSEMENTS	13.778	3.222		Total (Y)=20
	AVERAGE PERCENTAGE	81.047	18.953		74.074

Table 20: Lawshe test results for Domain B

ITEM	ELEMENT	Endorsement of statement		CVR	RETAIN (Y/N) (Reject if CVR < 0.49)
		Essential	Not essential or not necessary		
B1	**I find it easy to identify with my instructor.**	11	6	0.294	N
B2	**I can easily identify with my simulator partner.**	8	9	-0.058	N
B3	**I work well with others during simulator training exercises.**	9	8	0.058	N
B4	**Instructors communicate their expectations effectively.**	11	6	0.294	N
B5	**I learn better when I work as a member of the crew.**	17	0	1.000	Y
B6	**I am always at ease when interacting with my flight instructor.**	9	8	0.058	N
B7	**I always find my simulator partner prepared for training.**	8	9	-0.058	N
B8	**I trust my simulator partner.**	5	12	-0.411	N
B9	**I am confident that my instructor will be fair.**	7	10	-0.176	N
B10	**I operate well as a crew member in the simulator.**	15	2	0.764	Y
B11	**My instructor is willing to listen.**	14	3	0.647	Y
B12	**I communicate well with my simulator partner.**	14	3	0.647	Y
B13	**I feel secure in the decisions made by my simulator partner.**	9	8	0.058	N
B14	**I make good decisions with my partner in the simulator.**	6	11	0.294	N
B15	**I find that decision-making with my simulator partner is equitable.**	9	8	0.058	N
B16	**I am motivated by my instructor.**	5	12	-0.411	N
B17	**When training for this aircraft, I feel that I am part of a team.**	12	5	0.411	N

Table 20: Continued

ITEM	ELEMENT	Endorsement of statement		CVR	RETAIN (Y/N) (Reject if CVR < 0.49)
		Essential	Not essential or not necessary		
B18	**The instructors on this aircraft are committed.**	13	4	0.529	Y
B19	**Instructors are similar in how they teach pilots to fly this aircraft.**	16	1	0.882	Y
B20	**I am always paired with someone who is committed to performing well.**	10	6	0.176	N
B21	**I enjoy being evaluated as a member of a crew.**	3	14	-0.647	N
B22	**Instructors on this fleet follow company policy.**	9	8	0.058	N
B23	**The instructors on this aircraft avoid overloading pilots with unnecessary information.**	14	3	0.647	Y
B24	**I always bond well with my simulator partner.**	9	8	0.058	N
B25	**Decisions made in flight simulator training exercises are team-based.**	5	12	-0.411	N
B26	**The instructors on this aircraft are friendly.**	8	9	-0.058	N
B27	**I get sufficient feedback on my flight training performance.**	6	11	-0.294	N
	AVERAGE NUMBER OF ENDORSEMENTS	9.703	7.259		Total (Y)=7
	AVERAGE PERCENTAGE	57.076	42.924		25.926

Table 21: Lawshe test results for Domain C

ITEM	ELEMENT	Endorsement of statement		CVR	RETAIN (Y/N) (Reject if CVR < 0.49)
		Essential	Not essential or not necessary		
C1	**Pilots are in direct control of the training outcome.**	17	0	1.000	Y
C2	**A good training session on this aircraft is a result of the trainee's actions.**	11	6	0.294	N
C3	**Evaluation of my flight training is objective.**	11	6	0.294	N
C4	**Adequate preparation improves flight training performance.**	16	1	0.882	Y
C5	**I am always on time for a flight training session.**	14	3	0.647	Y
C6	**I co-operate well when training in a simulator.**	13	4	0.529	Y
C7	**I never feel rushed in the flight simulator.**	12	5	0.411	N
C8	**I easily express my opinion during flight training.**	5	12	-0.411	N
C9	**I prepare sufficiently for training on this aircraft.**	11	6	0.294	N
C10	**After flight training, I feel a sense of mastery.**	16	1	0.882	Y
C11	**I enjoy learning about this aircraft.**	7	10	-0.176	N
C12	**Simulator training affects behaviour on the actual aircraft.**	10	7	0.176	N
C13	**I get along well with my flight simulator partners.**	11	6	0.294	N
C14	**I found my transition to advanced automated aircraft easy.**	5	12	-0.411	N
C15	**I believe that if pilots do well in training, overall flight safety improves.**	10	7	0.176	N
C16	**I am happy with simulator training on this aircraft.**	11	6	0.294	N
C17	**I aim to do better at my next flight simulator training session by learning from my mistakes.**	11	6	0.294	N

Table 21: Continued

ITEM	ELEMENT	Endorsement of statement		CVR	RETAIN (Y/N) (Reject if CVR < 0.49)
		Essential	Not essential or not necessary		
C18	I have a positive relationship with my colleagues.	14	3	0.647	Y
C19	The workload between trainees is balanced during a flight simulator training session.	9	8	0.058	N
C20	Pilots are judged as members of a team when they train in the flight simulator.	7	10	-0.176	N
C21	I feel rewarded for the amount of work I put into flight training.	10	7	0.176	N
C22	The more work I put into my preparation for training on this aircraft, the better I will perform.	11	6	0.294	N
C23	Pilots who are prepared have no problems training for this aircraft.	14	3	0.647	Y
C24	It is essential that pilots prepare adequately to pass a rating on this aircraft.	11	6	0.294	N
C25	I am in control of the outcome of my flight training on this aircraft.	16	1	0.882	Y
C26	I enjoy studying the technical aspects of the aircraft.	15	2	0.764	Y
C27	I always learn something new after undergoing training on this aircraft.	11	6	0.294	N
C28	I focus on the pertinent and relevant topics when learning about this aircraft.	12	5	0.411	N
C29	I reflect on my learning after a flight training experience.	14	3	0.647	Y
C30	I look for additional information so as to gain a deeper understanding of this aircraft's systems.	16	1	0.882	Y
C31	I know where to find specific information for this aircraft.	11	6	0.294	N
C32	It is important to know more than just what is required to pass.	16	1	0.882	Y
C33	I aim to gain a deeper understanding of this aircraft.	14	3	0.647	Y
C34	I learn more than is required of me from the company.	6	11	-0.294	N

Table 21: Continued

ITEM	ELEMENT	Endorsement of statement		CVR	RETAIN (Y/N) (Reject if CVR < 0.49)
		Essential	Not essential or not necessary		
C35	**I find the training on this aircraft easy.**	9	8	0.058	N
C36	**I do well in training for this aircraft.**	11	6	0.294	N
C37	**I look forward to my next flight training session.**	10	7	0.176	N
C38	**I sleep well the night before training on this aircraft.**	10	7	0.176	N
C39	**An appropriate level of stress helps me perform well in flight training for this aircraft.**	6	11	-0.294	N
C40	**I'm comfortable undergoing training for this aircraft.**	14	3	0.647	Y
C41	**I can control my anxiety so as to perform well in training.**	13	4	0.529	Y
C42	**I enjoy spending extra time flight training.**	11	6	0.294	N
C43	**I am motivated to learn more about this aircraft.**				
C44	**I am happy to be subjected to regular flight checks.**	12	5	0.411	N
C45	**I enjoy route training on this aircraft.**	9	8	0.058	N
C46	**I enjoy simulator training for this aircraft.**	3	14	-0.064	N
C47	**If my simulator partner is having a bad day, I am not affected.**	8	9	-0.058	N
C48	**I create a relaxed atmosphere in the flight simulator.**	11	6	0.0294	N
C49	**The length of time spent simulator training is appropriate for this aircraft.**	5	12	-0411	N
C50	**I enjoy the free play flight simulator time on this aircraft.**	11	6	0.0294	N
C51	**I aim to gain a deeper understanding of this aircraft.**	6	11	-0294	N
C52	**I learn more than the company requires me to.**	12	5	0.411	N
	AVERAGE NUMBER OF ENDORSEMENTS	11.038	5.961		Total (Y)=15
	AVERAGE PERCENTAGE	64.93	35.07		28.846

The hypothesised constructed consists of three separate dimensions at an organisational (airline), group (instructor-trainee team) and individual (trainee) level of analysis. Thus, the statements were clustered accordingly. Tables 19 to 21 show the level of expert endorsement of each item within each dimension of the hypothesised construct. More importantly, the last columns in each table report on whether the item was retained or discarded, based on its content validity ratio.

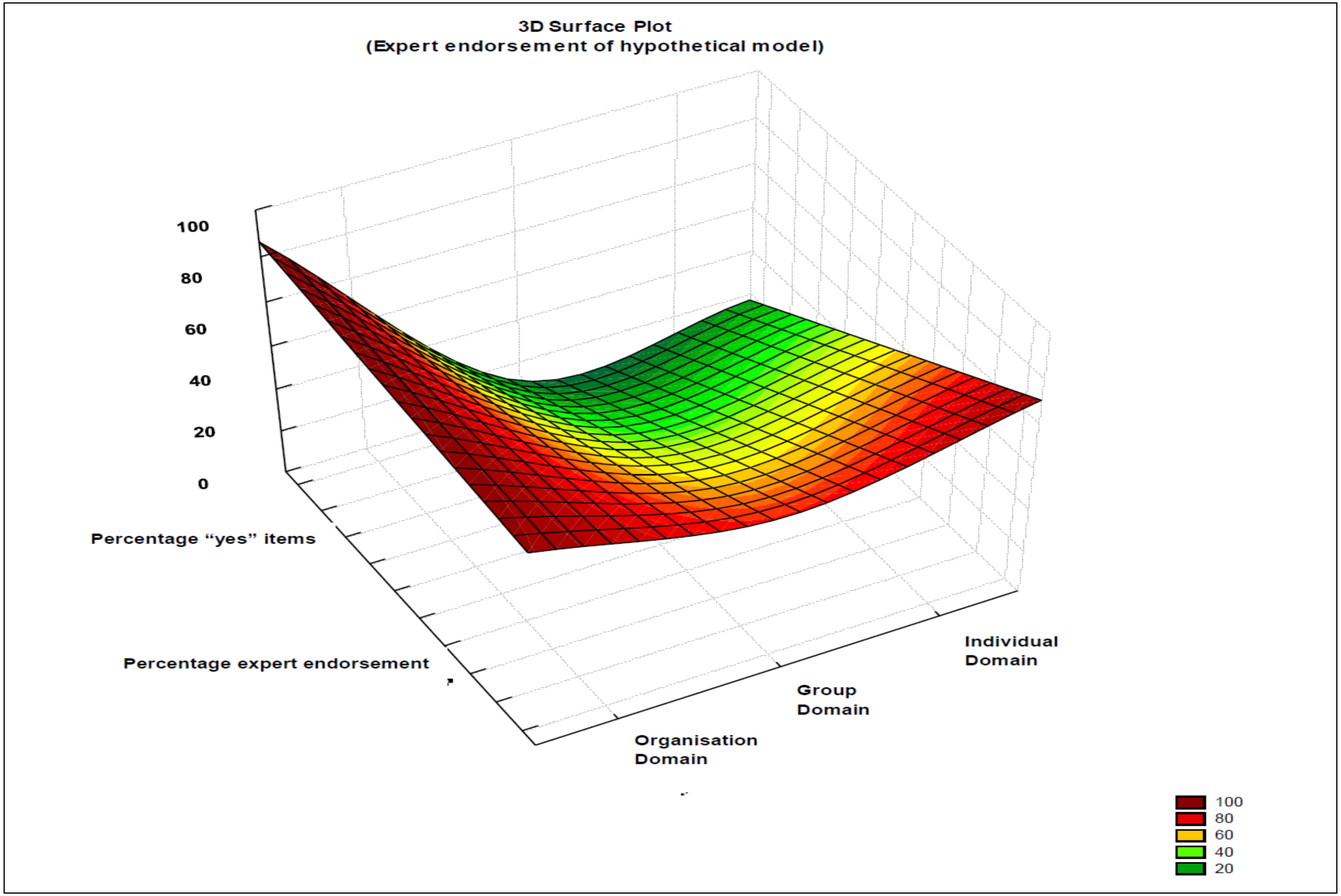
4.14 ITEM RETENTION RESULTING FROM THE APPLICATION OF LAWSHE'S TECHNIQUE

The computation achieved from the application of the Lawshe method resulted in the retention of 42 items. It appears that, out of the three dimensions of the hypothetical construct, Domain A, which assesses the organisational level of analysis (airline), was by far the most endorsed section, achieving an 81.047% proportion of expert endorsement and a 74.074% item retention level. Domain B, the group level of analysis of the hypothetical construct (the instructor-trainee dimension) received middling support from the panel of experts, with an endorsement proportion of 57.076% and an item retention level of 25.926%. Domain C, which assessed the trainee at the individual level of analysis, received a slightly higher level of support from the panel of experts, with an overall 64.93% acceptance of items, but only a 28.846% item retention level.

Additional clarity regarding the level of endorsement of items was mapped in a surface plot (see Figure 22). Myers, Montgomery and Cook (2009) refer to this kind of empirical evidence as a response surface model. The response given by the panel of 17 experts plotted on a three-dimensional surface suggests that items operationalizing the construct at a macro (organisational) level received far more support than the other two levels or dimensions. Red peaks suggest more support, whilst green troughs imply support to a lesser degree.

The 42 items extracted from the expert survey are considered very robust, due to the stringent criteria of the Lawshe method (Streiner, 2003). The next phase of scale development required an assessment of the authenticity of the data obtained from the item retention method followed.

Figure 22: Subject matter expert response surface model



4.15 ASSESSMENT OF INTER-RATER BIAS

Figure 22 usefully depicts the expert support for the 42 items that were retained. However, a level of inter-rater bias could not be eliminated and may have affected the analysis. Cochran's Q statistic was consequently calculated using the software package Statistica 7 to examine this possibility further. A matrix was produced in a one-way frequency table. A judge was given a score of 1 if he or she endorsed the proposed item, or conversely a score of 0 when the opposite was true (Table 22 provides a summary of this data). Therefore, a dichotomous variable was measured several times across differing conditions. According to Karlsson (2008), and Landis and Koch (1977), Cochran's Q test is an appropriate measure to determine whether the marginal probability of a positive response (that is, 1) is unchanged across the panel of judges. Cochran's Q test produced a very small P value (Q [16] = 201.3697, $p < 0.001$). Thus providing sufficient empirical evidence to conclude that the cohort of 42 essential or endorsed statements retained was of statistical importance.

Table 22: Summary of expert endorsement from Cochran's Q test

Expert	Sum	Percentage of 0s	Percentage of 1s
1	57	46.22	53.78
2	75	29.24	70.76
3	67	36.79	63.21
4	68	35.84	64.16
5	50	52.83	47.17
6	82	22.64	77.36
7	73	31.13	68.87
8	83	21.69	78.31
9	87	17.92	82.08
10	51	51.88	48.12
11	71	33.01	66.99
12	81	23.58	76.42
13	80	24.52	75.48
14	33	68.86	31.14
15	85	19.81	80.19
16	93	12.26	87.74
17	72	32.07	67.93
Mean	71.059	32.958	67.042

4.15.1 Final item retention

Based on the aforementioned analyses and further commentary from the group of subject matter experts with regards to the clarity and comprehensiveness of each retained item, Table 23 was produced after minor adjustments on selected statements. The final large survey item cohort is therefore found in the last column of Table 23.

Table 23: Comparison of items retained after applying Lawshe's method

Item	Retained statement based on Lawshe's method	Adjusted final large survey item
1	Pilot training at my airline is in line with company goals.	Training at my airline is in line with company goals.
2	My company's training produces world-class pilots.	My company's training produces world-class pilots.
3	I know what my company's training goals are.	I know what my company's training goals are.
4	My company has talented people managing airline pilots' training.	My company has talented people in training.
5	Pilot training at this company is professional.	Training on this aircraft is professional.
6	Management follows the regulator rules appropriately.	Management follows the rules and regulations appropriately.
7	Pilot training on this aircraft is well organised at this company.	Training on this aircraft is well organised.
8	I understand what the company expects of me when I am in training.	I understand what the company expects of me when training.
9	Training at my airline produces safe pilots.	Training at my airline produces safe pilots.
10	This airline gives its pilots an appropriate amount of preparation work before training.	The airline gives its pilots an appropriate amount of preparation work for training.
11	It is easy for me to appeal for assistance if I encounter a training problem at this airline.	If I had to experience a problem in training, it's easy for me to appeal.
12	There is sufficient training guidance from the company.	There is sufficient training guidance from the company.
13	The standard operating procedures (SOPs) for learning to fly this aircraft are adequate.	The standard operating procedures (SOPs) for learning to fly this aircraft is adequate.
14	The company provided me with sufficient time to prepare for training on this aircraft.	I'm given sufficient time to prepare for training on this aircraft.
15	The simulators my company uses to train its pilots are in good condition.	The simulators my company trains its pilots in are in good condition.
16	The airline is very supportive of its pilots' learning requirements for this aircraft.	The airline is very supportive of its pilots' learning requirements for this aircraft.
17	My company's culture supports training for new technology aircraft.	My company's culture supports training for new technology aircraft.
18	There is sufficient feedback about my training on this aircraft.	There is sufficient feedback about my training on this aircraft.
19	Pilot training at my airline follows civil aviation requirements.	Training is in line with civil aviation regulations.
20	My company uses only current training material.	My company uses only current training material.

Table 23: Continued

Item	Retained expert survey statement	Adjusted final large survey item
21	I learn better when I work as a member of the crew.	I learn better when I work as a member of the crew.
22	I operate well as a crewmember in the simulator.	I operate well as a crewmember in the simulator.
23	My instructor is willing to listen.	My instructor is willing to listen.
24	I communicate well with my simulator partner.	I tend to communicate well with my simulator partner.
25	The instructors on this aircraft are committed.	The instructor is committed.
26	Instructors are similar in how they teach pilots to fly this aircraft.	Instructors are very similar in how they teach pilots to fly this aircraft.
27	The instructors on this aircraft avoid overloading pilots with unnecessary information.	The instructors on this aircraft don’t overload us with information.
28	Pilots are in direct control of the training outcome.	Pilots are in direct control of the training outcome.
29	Adequate preparation improves flight training performance.	Preparation improves performance.
30	I am always on time for a flight training session.	I try never to be late for a training session.
31	I co-operate well when training in a simulator.	I co-operate when training in a simulator.
32	After flight training, I feel a sense of mastery.	After training I feel a sense of mastery.
33	I have a positive relationship with my colleagues.	I have a positive relationship with my colleagues.
34	Pilots who are prepared have no problems training for this aircraft.	Pilots who come prepared have no problems training for this aircraft.
35	I am in control of the outcome of my flight training on this aircraft.	I’m in control of the outcome of a training session.
36	I enjoy studying the technical aspects of the aircraft.	I enjoy studying the technical aspects of the aircraft.
37	I reflect on my learning after a flight training experience.	I reflect on my learning experience after a simulator session.
38	I look for additional information so as to gain a deeper understanding of this aircraft’s systems.	I read to understand so as to gain a deeper understanding of this aircraft’s systems.
39	It is important to know more than just what is required to pass.	It’s a good idea to know more than what is required.
40	I aim to gain a deeper understanding of this aircraft.	I aim to gain a deeper understanding of this aircraft.
41	I’m comfortable undergoing training for this aircraft.	I’m comfortable undergoing training for this aircraft.
42	I can control my anxiety so as to perform well in training.	I can control my anxiety so as to perform well in training.

4.15.2 Data collection

The self-administered survey (AATC-Q) was adopted for this part of the study. Three methods were used to distribute the large sample survey questionnaire to potential participants:

- Firstly, respondents were e-mailed a copy of the questionnaire, which they could answer, and then return to a specified e-mail address.
- Secondly, an electronic version of the questionnaire was hosted on the World Wide Web (see Appendix F). Additionally, the survey questionnaire web site was linked to the ALPA-SA home page and each potential respondent was requested to follow the link advertised.
- Finally, hardcopy questionnaire booklets were box-dropped in such a manner as to cover each pilot stratum. All responses to the hardcopy questionnaire were then subsequently recaptured on to the electronic version of the survey (World Wide Web).

4.16 DATA ANALYSIS

According to Cooper and Schindler (2003:87), "[d]ata analysis usually involves reducing accumulated data to a manageable size, developing summaries, looking for patterns, and applying statistical techniques".

In this section, the main approaches and techniques used to analyse the data that were collected are explained. One of the objectives of the study was to determine whether or not multiple variables contained in the measurement instrument could be reduced to a fundamental or latent factorial structure that may account for the majority of the variability found between respondents' replies.

In order to achieve the core research objective, the construct "perceptions of the advanced automated aircraft training climate" was operationalization and captured via an appropriate questionnaire as mentioned in Section 14.16.2.

4.16.1 Computerisation and coding of the data

Preparation of data requires concise editing, coding and statistical adjustment on the part of the researcher (Aaker *et al.*, 1995). The paper-based returns in this study required initial editing to identify omissions, ambiguities and errors. Answers that were deemed illegible or contained nonsensical responses were coded as “missing”. To ensure that this did not distort any interpretations of the data, the overall answers found in the returned questionnaires were reviewed. The paper-based returns were then recaptured electronically onto the web-based version of the survey to simplify data analysis. Coding the closed-ended questions from the web base was fairly straightforward, because the instrument made provision for response values and a column that was used for variable identification. The response values were then exported to a spread sheet and then entered into a computer software program. The Statistical Package for the Social Sciences (SPSS version 17) was employed to generate the statistical diagnostic information in most cases.

4.16.2 Statistical analyses

The purpose of conducting a statistical data analysis is to summarise univariate or multivariate data, to explore relationships between variables and to test the significance of these differences (Corston & Colman, 2003). The results obtained from the survey instrument were interpreted using appropriate statistical techniques for

- summary statistical descriptions;
- factor analysis;
- item analyses;
- reliability and homogeneity analysis;
- scale description;
- comparative analyses; and
- associational analyses.

The levels of measurement achieved at each stage also determined the choice of statistics used. The study treated the construct “perceptions of the automated aircraft training climate" as the dependent variable. In instances where the demographic variables of the sample frame were used to determine the effect of perceptions, these variables and situational categories then became the independent variables, and the hypothetical construct became the dependent variable, for example, the analysis of data would then indicate that more experienced airline pilots (the independent demographic variable) have a more favourable perception of the training climate (the dependent variable) than less experienced junior pilots have.

4.16.3 Analysis of compliance with specific assumptions

To assess compliance with the distribution requirements for factor analysis, Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy were used in this study. Morgan and Griego (1998:15) suggest that data are likely to factor well with a measure of sampling adequacy (MSA) of around "0.70". Table 24 shows the level of acceptability for the calculated measure of sampling adequacy according to Gravetter and Wallnau (2008).

Table 24: Acceptance levels for the measure of sampling adequacy

Acceptance	Measure of sampling adequacy (MSA)
Outstanding	0.90 to 1.0
Meritorious	0.80 to 0.89
Middling	0.70 to 0.79
Mediocre	0.60 to 0.69
Miserable	0.50 to 0.59
Unacceptable	Less than 0.50

As a general rule, for unequal sample sizes in social science research of this nature, Vermeulen (2009) strongly advocates computing Levene's test of homogeneity and Box's M-test for homoscedasticity. These diagnostic tests are administered to test for the assumption of equality of variance across groups. Such tests are a recommended requirement when conducting an analysis of variance or ANOVA when there is an assumption of the equality of covariance. In addition, favourable outcomes of these tests are sought for the parametric versions of a multivariate analysis of variance or MANOVA (Tabachnick & Fidell, 2007).

The Kolmogorov-Smirnov test (often called the K-S test) was used to analyse the normality of distributions and is generally regarded as the statistic of choice for such requirements in the behavioural sciences (Lilliefors, 1967). For instance, Field (2009) proposes that the K-S test be applied to determine whether a sample comes from a population with a specific distribution or can comply with a set of assumptions. The hypothesis regarding the distributional form (that is, the data following a specified pattern) is then rejected if the test statistic is greater than the critical value obtained from the SPSS-generated output table. Alternatively, Lilliefors (1967) and Pett *et al.* (2003), suggest conducting the *chi square goodness-of-fit* test to determine whether the observed frequency distribution of the respondents could reasonably have arisen from the expected sample frame distribution.

Both the K-S test and an analysis of the skewness and kurtosis of the data assisted the researcher in choosing between the two families of statistical methods, because choosing between a parametric and a non-parametric test can be difficult (Corston & Colman, 2003). Because one of the continued issues raised in survey research is the choice of statistics employed (Cohen & Lea, 2004), it was deemed important to critique the various methods available and to defend the final choices made, for achieving the goals in the present study.

Depending on the distribution pattern of the data received, appropriate parametric and non-parametric methods were considered at each analytical stage. According to Cohen and Lea (2004:222), a number of assumptions are generally made regarding the distribution of parametric variables:

- observations are independent;
- observations must be drawn from a normally distributed population;
- populations must have the same variances; and
- the means of these normal populations must be linear combinations of effects due to columns and/or rows.

Similarly, non-parametric assumptions may also have restricting requirements, such as, that

- observations must be independent; and/or
- the variable under study should have underlying continuity.

However, non-parametric testing tends to be far less restrictive than parametric procedures (Field, 2009). Stevens (1946) suggests that, instead of using actual measurements, the rank orders of measurements be used when conducting a non-parametric analysis. Depending on the situation, data was ranked from the highest to the lowest or vice versa (see Chapter 5).

Statistical tests fall into various categories of analyses, such as tests of differences and tests of relationships between groups or variables. Corston and Colman (2003) add that there is at least one non-parametric test that is the equivalent to any given parametric test (see Table 25). The categorisation and labels of some of these methods used in the final data analysis are summarised in Table 25.

Table 25: Comparison of statistical tests

	Parametric tests	Non-parametric tests
Differences between independent groups	• T-test for independent samples • ANOVA • MANOVA	• Chi square goodness of fit • The Kruskal-Wallis analysis of ranks • Mann-Whitney U • Non-parametric MANOVA
Differences between dependent groups	• T-test for dependent samples • Repeated measures ANOVA	• Wilcoxon's matched pairs test • Friedman's two-way analysis of variance
Relationships between variables	• Pearson's correlation coefficient • Probability regression analysis	• Spearman's Rho • Phi or Cramer's V • Kendall's Tau • Partial eta square • The chi square test

Source: Adapted from Cohen and Lea (2004), Field (2009) and Lilliefors (1967)

4.16.4 Descriptive statistics

Descriptive statistics were used to summarise the data. The essence of this statistical technique is to describe the sample and to calculate the mean, standard deviation, skewness and kurtosis of the sample scores (Corston & Colman, 2003; Gerbing & Anderson, 1988). An item analysis was then pursued to determine the initial item mean, item variance, standard deviation and item-scale correlation (Cooper & Schindler, 2003).

In order to analyse the distribution of each item as a percentage of respondents included in the different sub-dimensions, the descriptive statistical techniques mentioned were used where necessary. Any problems associated with the data that were collected (such as miscoded values or missing data) were discovered with the aid of summary statistics. Table 26 sets out the descriptive statistics that were applied in analysing the data obtained from the survey.

Table 26: Descriptive statistics

Summary statistic	Computation
Central tendency of variables	• Average or mean • Median • Mode
Measures of spread	• Variance • Standard deviation • Range • Inter-quartile range • Quartile deviation
Measures of shape	• Skewness • Kurtosis (platykurtic, leptokurtic, mesokurtic)

Source: Adapted from Cooper and Schindler (2003); Field (2009)

4.16.5 Factor analysis

Charles Spearman has been largely credited as the inventor of factor analysis (Cattell, 1987). Factor analysis is the preferred technique used to mathematically reduce a large amount of data into smaller more manageable clusters of related variables (Gerbing & Anderson, 1988). The method is commonly used in the behavioural sciences to uncover the latent dimensions when one is faced with a matrix of correlation coefficients (Cattell, 1987). Statistically clustering the common variables therefore informed the researcher of whether the instrument was a valid measure of the substantive constructs.

Two types of factor analysis were considered, namely exploratory factor analysis (which attempts to discover the nature of the underlying dimensions influencing a set of variables), and confirmatory factor analysis (which tests whether a set of variables is influenced by specific constructs in a predictive manner). Because the current study was an attempt to discover phenomena associated with a relatively unknown construct, an exploratory factor analysis was conducted on the dataset.

After receiving the questionnaires, respondents' answers were analysed by inserting the answers into a data matrix. The factor analysis used *heavy-duty* matrix algebra, because such a data matrix consists of as many rows as subjects (respondents), and as many columns as questionnaire items (Cohen & Lea, 2004; Pett *et al.*, 2003). In order to determine the interrelationships amongst these items, the data presented in the matrix took the form of Pearson product moment correlations or Pearson r (r_{xy}). According to Pett *et al.* (2003), the number corresponding to each row and column ranges from -1.00 to +1.00, where a negative value represents a negative correlation between the items, and a positive value represents the opposite. The subsequent meaning of relational strengths was then assessed in the context of the research topic.

Because one of the objectives in scale development is to determine the latent structure of a hypothetical construct, factor analysis was the analytical tool of choice to explain the variation and co-variation in a set of observed *variables* in terms of a set of unobserved *factors* (Field, 2009). Tabachnick and Fidell (2007) point out that reducing a complex array of data into statistically relevant correlates yields factors with some commonality. Sub-dimensions of the construct “perceptions of the advanced automated aircraft training climate” were then uncovered by using an exploratory factor analysis method. This technique is commonly employed when the exact number of factors that can accurately describe the construct of interest is unknown. In this case, there was no prior theory of the factorial structure of the construct that could be referred to. The basic aim of subjecting the data to an exploratory factor analysis was therefore to determine the relationship between observed, empirical evidence (survey results) and the latent factors (scale variables).

Field (2009) suggests that researchers use an appropriate statistical software package such as SPSS when conducting complex analyses. Many similar alternative software packages are also available in the market, such as Statistica. The program algorithm in many of the software packages calculates the interrelatedness between the factor, factors and/or other variables in the data space. This interrelatedness is then presented as a numerical value or correlation coefficient, referred to in exploratory factor analysis as a “loading” (Field, 2009). The loadings were used to find sub-constructs measuring the super-construct by rotating the loadings in order to find a pattern. The aim in this case was therefore to ascertain whether the variables of the measurement tool could be reduced (clustered) to yield an appropriate sub-structure.

4.16.6 Factor extraction

The retention of the correct number of factors has a significant bearing on the overall quality of a psychological scale (DeVellis, 2003; Gorsuch, 1997). Cohen and Lea (2004) contend that factors are generally extracted from a data set that represents the variance accounted for in each underlying factor. Two primary methods of extracting factors from the data space are available. Principal components analysis (PCA) or principal factor analysis (PFA) are the two methods of factor extraction often used in research similar to that in the current study. Schaap (2010) suggests that an exploratory factor analysis with principal factor analysis extraction is the preferred method. Additionally, according to Gorsuch (1997:534), “component analysis gives inflated loadings”. In the current study, therefore, a principal factor analysis was conducted on the combined questionnaire item response variables.

In principal factor analysis, the diagonal of the so-called big-R matrix is replaced by estimates of communalities, which may be the reason it is considered a more prudent method than principal component analysis (Schaap, 2010). According to Field (2009), the communality of a variable is the proportion of the variance that is produced by the common factors underlying the set of variables. However, the actual difference in results (the number and nature of the factors) obtained when contrasting a component's factoring extraction method can be small for data obtained in the social sciences (Warner, 2008).

When one needs to uncover the shared variance in a set of variables, Horn's (1965) parallel analysis, eigenvalues or Cattell's scree plot are techniques available to determine the number of factors that are to be retained. Gorsuch (1997) recommends that, if the sample is large enough (at least 300 cases), it should be divided into two sub-samples, and then each sub-sample should be subjected to a factor analysis. This comparison improves understanding of the extracted factors. For the purposes of this study, both parallel analysis and scree plots (which involve studying the slope of the plotted eigenvalues) were considered as a retention method. It was found in this study that the calculation of eigenvalues tended to produce many unnecessary factors that became difficult to examine without a scree plot or parallel analysis, because each eigenvalue is the percentage of total variance accounted for by a corresponding component.

The eigenvalue of each factor accounts for the number of variance units out of the total number of items that are being measured and that yield the approximate percentage of variance accounted for by a specific factor (Brown, Hendrix, Hedges & Smith, 2012). Furthermore, for a particular class of square matrices (A), it is possible to find vectors (eigenvectors, x) such that when said square matrix is multiplied by its associated eigenvector, the resultant product provides a scalar or constant value (λ), referred to in this case as the eigenvalue; that is, $A.x = x. \lambda$ (Brown *et al.*, 2012). Each element (item) of the square matrix would thus be associated with its own eigenvalue. In the current study, a major application of matrices was to represent such linear transformations, and therefore all the complex matrix algebra used in the study was conducted using commercially available computer software packages. In addition, the computed eigenvalues and eigenvectors provided an insight into the geometry of the transformations needed for factor analysis.

There are a number of criteria to determine the number of factors that should be retained (Gorsuch, 1983). Each of the following methods was considered based on both the advantages and disadvantages associated with the technique, and the requirements of the study:

- *Kaiser's criterion* – according to Kaiser's rule, components with eigenvalues under 1.0 should be rejected. However, this method tends to over-extract factors (Pett *et al.,* 2003).
- *"Variance explained" criteria* – this involves retaining enough factors to explain at least 90% or 80% of the variance in the data. This method is not recommended, given the level of subjectivity involved, and was thus discarded.
- *Cattell's scree test* – because a mathematical approach may lead to extracting factors of trivial importance (Gorsuch, 1983:167), Cattell (1987:16) suggests plotting the eigenvalues graphically. In this method, the components are plotted on the x-axis, with the corresponding eigenvalues plotted on the y-axis. The technique involves plotting the components as a diminishing series according to sizes and joining the points through the variables concerned. Where the number of factors ends due to certain error factors, a sharp break (or elbow) in the graph appears. This is why Cattell (1966:245) uses the analogy of "scree", which is a term that describes the broken rock fragments at the foot of a hill where it collects. Hayton, Allen and Scarpello (2004:192) point out that there is some subjectivity involved in determining the "sharp break" or when there are several "elbows" in the plot.
- *Comprehensibility* – the use of this method in *isolation* is not recommended as a scientific technique for answering the question of how many factors to retain. The non-mathematical nature of this process induces a fair amount of subjectivity when a researcher limits the number of factors to retain based on prior knowledge and comprehensibility.
- *Horn's Parallel Analysis (PA)* – a parallel analysis is one of the most robust and objective methods for retaining factors; however, very few computer programmes offer this solution, so the technique is rarely used. Because a parallel analysis

requires a Monte Carlo method to simulate mathematical systems (Glorfeld, 1995), Horn suggests comparing eigenvalues obtained from uncorrelated normal variables to the observed eigenvalues (such a comparison can only be achieved efficiently using a computational algorithm). Due to sampling error, sole reliance on Kaiser's criterion often overestimates the number of factors to retain; therefore Horn's method was considered in mitigation of this fundamental limitation (Hayton *et al.*, 2004). In order to determine the number of factors to extract without over- or underestimating the quantity, a modified version of Horn's parallel analysis was conducted in the current study, based on a Monte Carlo simulation in SPSS using the syntax developed by O'Connor (2000).

According to Sawilowsky (2003), a Monte Carlo simulation determines the properties of a phenomenon from repeated sets of random or permutated samples. Thousands of random or permutated samples can be easily generated using specialised computer-based algorithms. The choice between selecting randomised or permutated data sets is based on the level of robustness sort by the researcher, where, permutated parallel data sets are considered robust (more complex mathematical formulae) and therefore less susceptible to error. Also, the choice of selecting a method can be affected by the availability of the appropriate computer software programmes. However, in the current study, permutated sets of the original data were generated for comparison with the real data, as the option was available, and provides a more accurate solution. The steps followed in performing the analysis were the following, as recommended by Hayton *et al.* (2004):

- Permutated data sets were generated quickly, based on the same dimensions as that those analysed. This was possible using the syntax provided by O'Connor (2000), which produced the Monte Carlo type simulation.
- Next, eigenvalues from the permutated data correlation matrix were extracted, based on the principal axis option.
- The mean and 95th percentile of all eigenvalues generated from the permutated data sets resulted in a vector of the same size as the number of variables, and diminishing in value.

- Finally, the real data were compared in parallel to the permutations; in other words, eigenvalues from the real data and generated sets were compared. Statistically significant factors (in this case $p < 0.05$) were retained, based on the fact that they are greater than the eigenvalues from the permutated sets. The factors retained were therefore statistically significant and were due to more than chance.

Plotting the actual eigenvalues versus randomly generated or permutated eigenvalues can give a clearer picture of the solution. In this case, the graphical plots generated were also compared to Cattell's scree plot (discussed later in Section 5.2.2), providing a substantive level of confidence for factor retention.

For the purposes of the current study, Kaiser's criterion, Cattell's scree plot, Horn's parallel analysis and comprehensibility of factors were used in combination to determine the number of factors that should be extracted and subsequently retained.

4.16.7 Factor rotation

Once factors have been extracted, it is plausible that some variables have high loadings on one important factor and small loadings on all other factors, causing some confusion with the interpretability of the variables and their subsequent latent structures (Cohen & Lea, 2004; Field, 2009). Two methods of rotation were considered for this study, namely orthogonal rotations (varimax, quatimax, equamax) and oblique (oblimin, promax, direct quartimin). The goal of all rotation strategies is to obtain a clear pattern of loadings. A comprehensive discussion of the various sub-rotation methods is beyond the scope of the current study, and therefore only the rotation method selected for the present research is discussed.

Pett *et al*. (2003) advocate the use of Kaiser's normalisation method for factor rotation. This technique, which is the most common, and generally the default setting in commercial statistical software packages such as SPSS, was deemed appropriate for analysing the present data set (Kaiser, 1961). In a study similar in nature to the present one, Rogers, Monteiro and Nora (2008) found that Kaiser's normalisation decreased the standard errors of the loadings for the variables that had lower communalities and economised the correlations among oblique factors. However, some authors have justified the use of orthogonal rotations when employing a Kaiser normalisation, so as to enhance a better understanding of the latent factorial structure present within a data set, in particular, a varimax rotation is punted (Govindarajulu, 2001; Green & Salkind, 2008). This then leads to orthogonal factors and variables that are regarded as completely independent from each other (Glorfeld, 1995; Gravetter & Wallnau, 2008; Ho, 2006). Employing a varimax rotation with Kaiser's normalisation is still the most common technique in use. The method can result in lower eigenvalues; however, interpretability of the final factors is slightly superior. Nonetheless, Field (2009:643) warns that the choice of the type of rotation "depends on whether there is a good theoretical reason to suppose that the factors should be related or independent", and by observing the nature of variable clustering before rotation.

The behavioural sciences are considered an interdisciplinary genre and theoretical constructs are consequently related at some fundamental or root level. "…i[I]f one expects that the factors would be related among themselves, then an oblique rotation is appropriate" (Rogers, *et al*., 2008:261). Furthermore, employing specifically a promax oblique rotation "…maintains the same high loadings as the first orthogonal solution in a varimax factor analysis" (Rogers, *et al*., 2008:261). Therefore, an oblique rotation (promax) method as opposed to the orthogonal (varimax) method was deemed most applicable for a study of the present theoretical nature. The original premise then holds that the factors in a latent structure of the data are related.

The current study used a promax rotation with Kaiser's normalisation raised to a Kappa 4 (see also section 5.2.2 for a more in depth discussion), which included several rounds of exploratory factor analysis, until a definite structure of latent factors explaining the majority of the variability within the dependent construct was obtained. The term "promax" reflects that the new axes after rotation are free to take any position in the factor space (Corston & Colman, 2003:55). The method seeks a rotation (linear combination) where the degree of correlations among the factors is allowed, in general, to be relatively small because a pair of highly correlated factors should be interpreted as a single factor (Thurstone, 1947).

Due to the number of rotation and extraction methods available, some scholars have questioned the true objectivity of factor analysis (Hayton *et al*., 2004). Nonetheless, it is apparent that linear factor analysis continues to dominate the behavioural and psychological sciences as a method to assess dimensionality among a set of correlated or uncorrelated variables. To assess the latent structure of the data by determining which variables to retain, the current study considered factor loadings of 0.40 and above, as well as the cross loading of items on more than one factor, and the reliability and importance of each variable.

The information for interpreting a factor analysis was obtained from the summary table output produced by the SPSS software programme (version 17), and is reported in Section 5.2. The summary table (see Table 35) relates factor loadings, communalities, eigenvalues, and the cumulative percentages of variance.

4.16.8 Reliability

In general, a test's reliability is examined by estimating the amount of error associated with its scores: "One of the central tenets of classical test theory is that scales should have a high degree of *internal consistency*, as evidenced by Cronbach's alpha (α), the mean inter-item correlation, and a strong first component...is used in establishing the reliability of the scale" (Streiner, 2003:217). Cronbach's coefficient alpha is the most commonly used statistic in determining the *reliability* of a scale (Field, 2009:674). Reliability in the current study implies that the measure that was developed consistently reflected the construct that it purported to measure. The alpha statistic is mathematically defined in the reliability formula (Field, 2009; Streiner, 2003; Zeller & Carmines, 1980):

$$\alpha = \frac{N}{N-1}\left(1 - \frac{\sum_{i=1}^{N} \sigma_{Y_i}^2}{\sigma_X^2}\right)$$

Where:

α = Cronbach's coefficient alpha;
N = The number of items
$\sigma^2 X$ = The variance of the observed total item scores
$\sigma^2 Yi$ = The variance of component i
i = Component i

The above equation clearly shows that α increases when the correlations between items increase. In theory, Cronbach's coefficient alpha should not be considered a statistical test *per se*, but should instead be related to a coefficient of reliability or indication of consistency (Zeller & Carmines, 1980).

In essence, by determining the reliability of an instrument, it is established just how well the items reflect the same construct and may consistently produce similarity in results (Saunders *et al*., 2007). The coefficient was used in the study to provide a mathematical value of how well the set of items measured the latent construct. The alpha statistic is a good measure of reliability because it relates directly to the average co-variance of all the items and is inversely proportional to the sum of all the item variances and co-variances (DeVellis, 2003). This therefore also provided the statistical level at which the items in the current research scale (or sub-scale) correlated with each other. "The high correlation tells us that there is similarity (or homogeneity) among the items" (Cooper & Schindler, 2003:239). This is an important concept that was a basic requirement in the development of the current psychological and behavioural measurement instrument.

An examination of the literature to determine the best value for Cronbach's coefficient alpha revealed a lack of complete clarity and also presented some confounding arguments (see Table 27). Most authors however, are comfortable in converging on a cut-off value of 0.70 as a reliable measure of perception (Clark & Watson, 1995:310; Cooper & Schindler, 2003:629; DeVellis, 2003:28; Field, 2009:675). In addition it can be argued that values in excess of 0.90 do not necessarily demonstrate good reliability, but rather point to the redundancy of items in a scale (Streiner, 2003).

The raw Cronbach's coefficient alpha values were compared to the values of the alpha value pertaining to the group of items in each sub-scale by deleting the item under analysis. An increase in the overall value of alpha indicates that the variable is neither reliable nor valid, and that it should be excluded from further examination (Cooper & Schindler, 2003; Streiner, 2003).

Table 27 was used as a guide in determining the level of item internal consistency for the developed measurement scale, as it provides a comparison of acceptable alpha values.

Table 27: A contrast of relevant Cronbach's coefficient alpha values

Source	Rationale
Cortina (1993:102)	α = 0.70: If the scale contains more than 14 items, or when the scale has at least two orthogonal dimensions with modest (0.30) inter-item correlations, 0.70 is a good alpha value for the test of reliability.
Field (2009:679)	α = 0.70 to 0.80: This shows that a questionnaire has good overall reliability; and 0.70 is needed for ability tests.
Netemeyer *et al*. (2003:102)	α = 0.70: 0.70 is widely advocated as an adequate alpha measure, and the statistic is directly related to the number of items and inter-item correlations.
Streiner (2003:102)	α = 0.70 to 0.90: An alpha higher than 0.90 may indicate redundancy rather than improved levels of scale reliability

4.16.9 Homogeneity

Within the framework of the theory of tests and measurements, homogeneity relates to the degree to which the items approximate a hierarchical scale (Krus, 2006). Alternatively, homogeneity can be used to test the inter-relatedness of each item with another and therefore it was used to determine the efficiency of the AATC-Q in measuring particular constructs (Oosterhof, 1976).

A comparison of the coefficients of reliability (r_{xx}) and homogeneity (h_{xx}) is provided in Table 28.

The formulae depicted in Table 28 show that both reliability and homogeneity are a function of the mean squares (MS), which represent the average squared deviations of an effect of interest around the grand mean. The mean inter-item correlations of the items in each factor were used to establish the unidimensionality or homogeneity of the scale (Clark & Watson, 1995). Inter-item correlations for the current study exceeded 0.20 and were thus regarded as acceptable in terms of the criteria advocated in the literature (De Vellis, 2003).

Table 28: Summary of reliability and homogeneity coefficients

Reliability	Homogeneity
Spearman-Brown's coefficient of reliability	Guttman's coefficient of reproducibility
Kuder-Ruchardson's formula K-R 20	Loevinger's (1957) coefficient of homogeneity
Cronbach's coefficient of reliability (alpha)	Cliff's coefficient of homogeneity
Hoyt-Jackson's coefficient of reliability $r_{xx} = \frac{MS_R - MS_I}{MS_I}$	Homogeneity as formulated by Krus and Blackman $h_{xx} = \frac{MS_R - MS_I}{MS_R^{max} - MS_I^{max}}$

Source: Adapted from Krus (2006)

4.16.10 Item discrimination analysis

After determining the appropriate items that contributed statistically to the latent factors of the measurement construct, it was necessary to examine the dispersion of the scores. A discriminant analysis provided evidence that the items in the developed scale were effective, in other words, the items differentiated adequately between high scores and low scores. In the study, item discrimination indexes were calculated by separating the item mean scores of the top half and lower half scores of responses from each item in the main interacting construct factors as suggested in Oosterhof (1976).

Furthermore, the upper bound scoring participants were separated from the lower bound scoring participants by appropriate dummy variables (1 or 0), and their response patterns were explicitly modelled based on the difference between the data classes. In discriminant analysis, according to Leech, Barrett and Morgan (2005:131), "one is trying to devise one or more predictive equations to maximally discriminate people in one group from those in another group". Statistically significant differences between discriminant item groups provided some evidence that the AATC-Q is a highly effective scale.

Due to the potential violations in the assumptions of multivariate normality in the collected data, poor accuracy of the estimates of the probability of correct classification was anticipated and expected. However, the method was still highly valid in achieving the study's scale development goals, because discriminant analysis is fairly robust to the assumptions of linearity, normality and equivalence of covariance across groups (Embretson & Reise, 2000). Nonetheless, both a matrix scatter plot and Box's M-test was used to determine the assumption of homogeneity of variance-covariance matrices in the data. Where the scatter plots were fairly equal, homogeneity of variance-covariance was assumed (Leech *et al*., 2005) and the power of the discriminant model was considered relatively stable.

4.16.11 Comparative statistics

To compare the relationship between the demographic dimensions of the respondents in the sample and the main and interacting constructs and sub-constructs, it was necessary to utilise various univariate and multivariate procedures (Field, 2005; 2009). The following measures of analysis of variance were selected as the basis for making inferences about the current data set.

Multivariate analysis of variance (MANOVA) was used to determine the main and interaction effects of categorical variables on multiple dependent interval variables. MANOVA makes use of one or more categorical independents as factor variables, with two or more dependent variables (Morgan & Griego, 1998). MANOVA tests the differences in the centroid (vector) of means of the multiple interval dependents, for various categories of the independent(s). After an overall F-test had shown significance, *post hoc* tests where then used to enable a more precise evaluation of differences between specific centroids. It has been suggested that the *post hoc* multiple comparison tests be performed separately for each dependent variable (Field, 2005:571-595). For example, the categorical subgroups (such as Boeing, Airbus, male, female) of the sample group were compared independently using the proposed *post hoc* procedures.

Because the data set did not meet the assumption of normality or homogeneity of variance-covariance (see Table 54 for more detail), the "non-parametric MANOVA with rank order data" was performed in the present study (Zwick, 1985:148-152). Box's M test was used to ascertain the homogeneity of the variance-covariance matrices (Anderson, 2001; Clark & Watson, 1995), because Zwick (1985) warns that the power of a significance test can be severely aggravated whenever sample sizes vary across cells. SPSS (version 17) was used to determine Box's statistic, and a violation of the assumption of homogeneity would then be accompanied with a low p-value, therefore this then further substantiated the employment of a non-parametric MANOVA.

To determine the significance of differences between categories (based on the multivariate Pillai-Bartlett trace, V), the degrees of freedom are initially computed by multiplying the number of dependent variables to the number of groupings minus one (Anderson, 2001; Zwick, 1985). Based on an appropriate alpha level, a chi-square table was consulted to locate the chi-square value that pertains to these degrees of freedom. Whenever the difference between the test statistic and critical chi-square value was exceeded, the test would be considered significant (Field, 2005; Zwick. 1985). Furthermore, the familiar Mann-Whitney non-parametric *post hoc* tests were applied to calculate the differences between the rank ordered means with only two categories or sub-groups (Morgan *et al*., 2007).

Due to the skewness of the distribution of the responses in this study, it was also decided to use the Mann-Whitney U test and Kruskall-Wallis test to compare the mean rank order scores of different groups (see Section 5.6 for detail). These tests are commonly referred to as distribution-free tests (Rencher, 2002; Rosenthal, 1994). As non-parametric methods, their applicability is much wider than the corresponding parametric methods. Because non-parametric methods rely on fewer assumptions, they are also more robust (Field, 2005; Stuart, Ord & Arnold, 1999). The final analyses employed non-parametric or distribution-free statistical tests, because these tests do not depend on any assumptions about the form of the sample population or the values of the population parameters, making the method less limiting.

The Mann-Whitney U (M-W) test is a non-parametric test used to assess whether two samples of observations come from the same distribution (Tabachnick & Fidell, 2007). This test was used because the assumptions of the t-test were violated, in that the dependent variable data set was non-normally distributed or ordinal in nature. The Mann-Whitney U test is only slightly less powerful then Student's t-test (Field, 2005; Morgan *et al.*, 2007). For the M-W test, Z-values are calculated that are used to "approximate" the statistical level of significance for the test (Winks, 2008:108). The method was found to be ideal for unequal and small sample sizes (Babbie, 2010).

The Kruskall-Wallis (K-W) test is a non-parametric test that can be applied to assess whether three or more independent samples of observations have the same distribution (Rencher, 2002). The Kruskall-Wallis test was used as an alternative to its parametric one-factor ANOVA counterpart because the ANOVA's normality assumptions were not completely met in the present data set; also the data were ordinal. The K-W test uses mean ranks to determine whether scores differ across groups and a chi-square distribution to estimate the statistical level of significance for the test (Field, 2005; Morgan *et al.*, 2007).

4.16.12 Associational statistics

When one explores the relationship between variables, one should quantify the degree of linear relationship between two variables at an ordinal or interval level of measurement (Embretson & Reise, 2000). For an ordinal level of measurement, Spearman's Rho can be used, and for an interval level of measurement, Pearson's Correlation Coefficient can be used, thereby generating a value of association between parametric variables (Field, 2009). Alternatively, Netemeyer *et al.* (2003) suggests the chi square test and Kendall's Tau-b statistic as a more robust determination of association when dealing with non-parametric data. These techniques (see Table 29) were kept in cognisance when attempting to understand the phenomena within the data set.

Correlational research is used to discover how the status on one variable tends to reflect the status on another (Babbie, 2010). Mainly non-parametric correlations were used in this study to predict the effect of one variable on another; and examine related events, conditions or the behaviour of the population sample. According to Morgan and Griego (1998), the predictor variable (independent) is believed to produce an outcome in the dependent or criterion variable. In order to establish the most appropriate correlation statistic to use in determining association, Table 29 was consulted for contrasting the available options.

Table 29: Correlation statistic guideline

Predictor variable	Criterion variable	Correlation to use
Interval (continuous)	Interval (continuous)	Pearson
Real dichotomy	Interval (continuous)	Point biserial
Artificial dichotomy	Interval (continuous)	Biserial
Real dichotomy	Real dichotomy	Phi
Artificial dichotomy	Artificial dichotomy	Tetra choric
Ranking	Ranking	Spearman's rho
Ranking	Ranking	Kendall's tau

Source: Adapted from Field (2009), Leech *et al.* (2005) and Stuart *et al.* (1999)

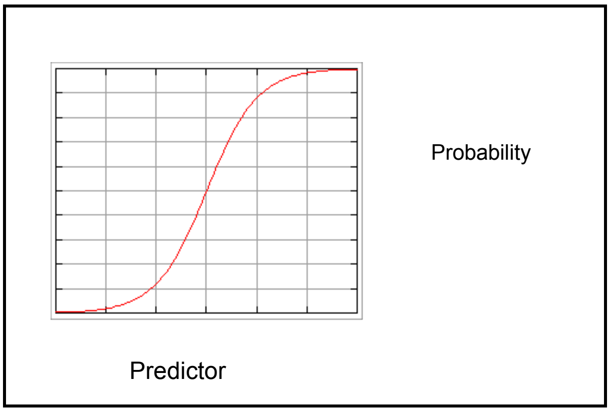
4.16.13 Logistic regression analysis

Logistic regression was the method of choice in developing a predictive model from the data set, because "[r]egression methods have become an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables" (Hosmer & Lemeshow, 2000:1). The probability of a binary outcome on a discrete variable was modelled from the most likely relationship between demographic covariates as a last step in the final statistical analysis of the observations.

"Logistic regression allows one to predict a discrete outcome such as group membership from a set of variables that may be continuous, discrete, dichotomous or a mix" (Tabachnick & Fidell, 2007:437). A stepwise method was adopted, where all predictors were initially placed in the logistic model and eliminated sequentially in subsequent steps of the analysis. Stepwise regression is used in the exploratory phase of research (Field, 2005). The options available in a logistic regression and justification for the final method selected is discussed in Section 5.10.

In the current study, a dichotomous dependent variable was constructed, based on the level of favourability perceived by the respondents. Similar to discriminant analysis, a dummy variable (1 or 0) was allocated to the dichotomy of perceiving a favourable or unfavourable training climate. The regression based on the logit is particularly useful in this case, as the distribution of responses on the dependent variable was non-linear, with one or more of the independent cases or co-variates. In other words, the probability that the criterion variable will have one outcome rather than another is based on a non-linear function of the best linear combination of the predictors (Tabachnick & Fidell, 2007). The logistic curve can be considered a generalised sigmoid or “S” curve (Govindarajulu, 2001; Morgan *et al.*, 2007). An “S” curve begins exponentially and thereafter begins to taper off (see Figure 23). The curve is based on a mathematical concept that has been widely used to model the natural life cycle of many phenomena (Cohen & Lea, 2004). For instance, in the case of the current study, the plotted logistic regression model clearly showed an initial exponential change in probability between the levels of perceived computer competence in the interaction effect between flying experience and computer ability, and a slowing down or tapering off later in the curve.

Figure 23: General shape of the common sigmoid curve used in logistic regression



Source: Adapted from Govindarajulu (2001)

The non-parametric nature of the empirical data that was collected in the current study meant that an alternative multiple regression analysis technique was required. According to Field (2009), the logistic function provides a value of probability between 0 and 1 based on the logit formula or curve (non-parametric), where logit (p) = ln (p/1-p), also referred to as the log odds (Tabachnick & Fidell, 2007). This is in contrast to the probit curve (parametric), which is based on a probability unit or normal distribution (Govindarajulu, 2001). The primary reason for using a logistic regression analysis to model the data in this study was that the method offers several distinct advantages (Tabachnick & Fidell, 2007):

- logistic regression is robust, in that the independent variables in the equation do not have to be normally distributed or display equal variances within groups;
- there is no assumption of linearity between the predictor and criterion variables;
- the outcome variable can be binary; and
- there is no requirement for continuous or interval independent variables.

However, Field (2005) points out that there are also some distinct disadvantages associated with the logistic regression technique. For instance, the method requires a higher number of data points to produce meaningful and stable results.

To determine how powerful the developed regression equation was at predicting the variable of interest (proportion of variance in the criterion variable associated with the predictor variable), a pseudo R^2 was computed, based on the methods of Cox and Snell's R-Square, Nagelkerke's R^2, and McFadden's (adjusted) R^2. A pseudo R^2 is computed to evaluate the goodness-of-fit of the logistic model (Tabachnick & Fidell, 2007). In general terms, a correlation coefficient can range between -1 and 1. However, in the case of the R^2 coefficient, the value computed in the current study ranged from 0 to 1 (because squaring the correlation between the predicted values and the actual values of the regression model produced a positive value). This value is referred to as the so-called pseudo R^2. According to Govindarajulu (2001), a high value of the pseudo R^2 indicates that there is a greater magnitude of the correlation between the predicted values and the actual values. Ho (2006) cautions, however, that when using different pseudo R^2s, it is possible that one may arrive at very conflicting results.

Cox and Snell's R^2 was calculated based on the following formula (Long, 1997):

$$R^2 = 1 - \left\{ \frac{L(M_{Intercept})}{L(M_{Full})} \right\}^{2/N}$$

Where L(M) is the conditional probability of the dependent variable, given the independent variables (the model *intercept* contains no predictors or independent variables). When there are N observations, L(M) is the product of N probabilities. The formula shows clearly that even if the regression model were a perfect fit, the R^2 value can never attain a value of 1 (Field, 2005). Cox and Snell's R^2 indicated the improvement from the null model (intercept only) to the derived or fitted model.

Nagelkerke's R^2 was calculated based on the following formula (Long, 1997):

$$R^2 = \frac{1 - \left\{ \frac{L(M_{Intercept})}{L(M_{Full})} \right\}^{2/N}}{1 - L(M_{Intercept})^{2/N}}$$

Like Cox and Snell's R^2, Nagelkerke's R^2 provides a value which indicates the improvement of the full model from the null or intercept only model. However, the formula indicates clearly that the range of the R^2 value can achieve 1 in a prediction model with a perfect fit.

McFadden's (adjusted) R^2 was calculated based on the following formula (Long, 1997):

$$R_{adj}{}^2 = 1 - \frac{\ln \hat{L}(M_{Full}) - K}{\ln \hat{L}(M_{Intercept})}$$

In the above formula, L-hat refers to the estimated likelihood. The adjusted formula is useful in that it indicates whether particular predictors add value to the model or not. The model is penalised by a reduced R^2 value when there may be too many ineffective predictors (K). The formula clearly shows that a negative R^2 is possible when using McFadden's adjusted method.

Several significance tests are available to determine the inclusion or exclusion of co-variates from a logistic regression model (Hosmer & Lemeshow, 2000). The Wald test was used to test the statistical significance of each of the coefficients in the regression model. A Z-score (Z = coefficient [B]/SE) was calculated. The hypothesis of inclusion or exclusion of the coefficients was thus based on the subsequent chi-square fit (Tabachnick & Fidell, 2007). For smaller sample sizes, Agresti (1996) suggests the likelihood-ratio test. Because a backward stepwise elimination was followed in building the final model, the likelihood-ratio test statistic used in the current study was based on the following formula:

$$-2\log\left(\frac{L_0}{L_1}\right) = -2[\log(L_0) - \log(L_1)] = -2(L_0 - L_1)$$

The above equation shows that the maximised value of the likelihood function for the full model (L_1) was compared to the maximized value of the likelihood function for the simpler or null model (L_0), associated with a chi-square goodness-of-fit.

The Hosmer-Lemeshow goodness of fit test was applied to determine whether the model prediction did not differ significantly from the observed number of subjects in each group. It was desirable to achieve a non-significant chi-square test statistic in a case such as this, as recommended by Agresti (1996). A good model would be effective in categorising most successful subjects into an upper level and placing failures into a lower level (Hosmer & Lemeshow, 2000).

The change in odds (odds ratio) based on a unit change in the predictor indicated the impact of each predictor on the regression model. The odds ratio and the other tests mentioned which were conducted on the final logistic regression model are discussed further in Section 5.10.

4.16.14 Practical significance and effect size

When a significant result is reported from the research, it is accompanied by a p-value and confidence can be established in the assumption that the results are not simply due to chance (Muijs, 2004). However, a p-value has more meaning when it is accompanied by an effect size value. Cohen and Lea (2004) are critical of studies that assess significance without an accompanying practical or effect size report. For instance, if a study looks for a 95% confidence interval, a p-value of less than 0.05 implies that there is less than a 5% probability that the result may occur by chance. However, this would not indicate how large the significance actually is.

The goal of practical significance computations of research results is an interpretation to present significant conclusions, which may be more meaningful to a non-statistician (Cohen, 1992). In other words, statistical significance shows that the results are unlikely to have occurred by chance, whereas practically significant and effect size results are more “meaningful in the real world” (Ellis, 2010:15). It is important to note that in this context effect size does not refer to cause and effect relationships between variables, but merely provides a value that quantifies the practical significance of findings (Rosenthal, Rosnow & Rubin, 2000).

In many cases, it is necessary to know whether a relationship between two variables is practically significant – for example, between pilots’ level of education and their perceptions of advanced automation training climate. The statistical significance of such relationships can be determined by using the correlation coefficients (r). In this case, the effect size was determined by using the absolute value of r and relating it to the cut-off points for practical significance recommended by Cohen (1988), where

- $r = 0.10$ suggests a small effect;
- $r = 0.30$ suggests a medium effect; and
- $r = 0.50$ suggests a large effect.

To assess the significance of the z-statistic of the Mann-Whitney test the coefficient ‘r’ was computed by using the conversion formula, $r = z/\sqrt{(N)}$ suggested by Field (2005) and Morgan *et al*. (2007).

Pett *et al*. (2003) recommend the calculation and use of the partial eta square (η^2) to determine the effect sizes or strength of relationship between demographic variables and the construct of interest. The results of the η^2 provide a value that quantifies the practical significance of the findings (Cohen, 1988). Where MANOVAs and ANOVAs are implemented and statistically significant main and interaction effects are found, the partial eta squared is calculated to determine the practical effect size. Partial eta squared (η^2) is the proportion of the effect, plus the error variance attributable to the effect. In the current study, Field’s (2009) formula, $\eta^2 = (SS_{effect})/(SS_{effect} + SS_{error})$, was used to determine partial eta squared.

According to Cohen’s (1988) effect size criteria, the following cut-off points normally apply if partial eta squared is to be of practical significance:

- $\eta^2 = 0.01$ suggests a small effect;
- $\eta^2 = 0.06$ suggests a medium effect; and
- $\eta^2 = 0.14$ suggests a large effect.

4.17 RESEARCH ETHICS

An area of concern for many researchers conducting empirical studies using primary data in the social sciences is gaining access to participants (Saunders *et al*., 2007). To overcome this problem, permission was obtained from ALPA-SA, which gave its consent for the researcher to access its database of over 1 000 airline pilots. Airline management at the various organisations also endorsed the project. In addition, all participants were asked to acknowledge a consent form prior to commencing with the survey (see Appendix D). In accessing these participants, a quality scientific inquiry was conducted, adhering strictly to the moral and ethical research principles applicable at the University of Pretoria.

The *quality* of a research design is directly related to the *ethical* standing of the final research (Rosenthal, 1994:127). Researchers in the social and psychological sciences face many dilemmas that may have an impact on the morality of their studies. Some important issues in ethics that can become a problem if these issues are not avoided are the following (Rosenthal, 1994):

- *hyper-claiming* – claiming that the research will achieve specific goals and objectives that it cannot achieve;
- *causation* – claiming that there is a causal link between variables when there is actually none;
- *data dropping* – analysing data that never existed or removing data that conflicts with the researcher's hypotheses; and
- *questionable generalisations* – failing to pay careful attention to not inferring findings on a population without sufficient empirical evidence or an adequate sampling technique.

An outline of the principles involved at each stage of the research process to ensure that the study was conducted in an ethical manner is provided in Table 30. The columns show the stage of the research at which a particular principle was applied and thereafter what techniques were used to mitigate any adverse morality issues.

Table 30: Ethical issues considered in the research process

Stage of research	Possible ethical issue	Specific issues addressed
Exploration	Confidentiality	Sponsor non-disclosure.
Research proposal	Informed consent	Participants' and sponsors' right to quality research.
Research design	Informed consent and confidentiality	Deception of respondents. The right to privacy. Immoral coercion from parties with ulterior agendas.
Data collection/preparation	Confidentiality	Participant privacy issues. Data exploitation.
Data analysis/reporting	Confidentiality and Data completeness/integrity	Confidentiality of participants. Censoring of results. Meta-analysis.

Source: Adapted from APA (1994), Cooper and Schindler (2003), and Rosenthal (1994)

The issues highlighted in Table 30 were dealt with on a case-by-case basis during the writing and distribution of the final completed questionnaire.

According to Cooper and Schindler (2003:120), "[e]thics are norms or standards of behaviour that guide moral choices about our behaviour and relationships with others" – hence, any research conducted in the field of psychology affecting human subjects requires consent from the organisation responsible. The American Psychological Association's (APA) guidelines were strictly adhered to in order to maintain ethical standards required by the University of Pretoria. These principles, according to the APA (2002:3-5) are

- *beneficence and non-maleficence* – psychologists should maximise the benefits of participants and minimise any harm that may result from their research;
- *fidelity and responsibility* – scientists must establish lines of trust between themselves and participants, and must ensure the highest standards of professionalism by maintaining objectivity;
- *integrity* – scientists must promote honesty, accuracy and the truthfulness of their research;
- *justice* – all persons participating should be entitled to access the research that is being conducted; any unjust practices must be prevented by guarding against potential biases; and
- *respect for people's rights and dignity* – researchers must adhere to the requirements of participants' right to privacy, confidentiality and self-determination at all times.

The final instrument was accompanied by a cover letter introducing the study. The letter assured respondents who volunteered to take part in the study of the confidentiality of their responses and their anonymity. The current study endeavoured to maintain academic objectivity at all times by following a structured design and methodology, and complying with the ethical requirements for research of this kind. Annexure D contains a draft of the informed consent form that was used.

4.18 SUMMARY

The chapter focused on the methodologies and statistical applications required to design and construct a valid and reliable psychometric instrument for the advanced aircraft industry. The present research was completed by means of abductive, inductive and deductive reasoning processes grounded in prior theory. The study design consisted of an empirical quantitative approach based on a positivist paradigm, which resulted in the development of a new measurement tool in the advanced aircraft training environment.

The chapter also discussed the population, method of sampling, the design and layout of the questionnaire, the type of questionnaire used, the design of the questionnaire items, as well as the correlations, factor analysis, comparative, associational and regression modelling technique used in the study. Practical and effect sizes were discussed with regard to reporting any tests of significance.

The statistics used in the research were discussed rather in detail, because they form the foundation for the reporting of results and recommendations set out in the subsequent chapters.

The following general quantitative steps in the second phase of the research were undertaken to meet the research objectives:

- Step 1: Determine the content validity of each item’s relationship with the construct and sub-constructs. This was calculated using,
    - Lawshe’s (1975) method, and
    - Cochran’s Q statistic.

- Step 2: Explore the data using a factor analytic technique;

- Step 3: Applying an appropriate rotation method to extract an optimum number of factors;

- Step 4: Analyse clusters of items;

- Step 5: Determine the level of reliability and homogeneity present;

- Step 6: Summarise the data;

- Step 7: Determine the distribution and normality status of the data set; and

- Step 8: Explore possible relationships and phenomena of the latent structure by means of the following non-parametric statistical procedures:
    - the non-parametric MANOVA;
    - the Kruskall-Wallis test;
    - the Mann-Whitney U-test;
    - Spearman’s rank order correlation and Kendall’s tau; and
    - Logistic regression.

CHAPTER FIVE:

RESULTS

5.1 INTRODUCTION

The primary purpose of the study was to develop a valid and reliable instrument to measure airline pilots' perceptions of the training climate associated with advanced automated aircraft. The fundamental goal was to develop a questionnaire by operationalizing a hitherto unobserved hypothetical construct (perceptions of the advanced automated aircraft training climate) based on the three hypothesised levels of analysis (the person, the group and the organisation) that conceptualised the construct. In addition, the purpose of the research was to explore the relationships between the data and the characteristics of the data further in terms of the latent constructs that emerged.

The earlier part of the discussion provided a theoretical overview of the background in which the hypothetical measurement construct was devised and developed. Chapters 5 and 6 report on the results, interpret and discuss them.

Factor analysis was the statistical method used to provide scale descriptors, and it played a pivotal role in the development of an appropriate psychological measurement tool. Furthermore, the results of an analysis of the items forming the scale and the reliability of the constructs are reported on and discussed. To enhance the strategic statistical exploration of the phenomena present in the data set, initially, basic non-parametric comparative and associational analyses were completed. Subsequently, more in-depth and complex analyses, such as non-parametric MANOVAs and a stepwise logistic regression, were conducted to round off a thorough examination of the data set.

5.2 FACTOR ANALYSIS

Developing a scale for any psychological measurement instrument generally entails reducing the attribute space of a large number of variables into correlated factors or dimensions (Clason, & Dormody, 2001). Thus, it was assumed that the explanation for any correlations between the many variables in the dataset could be obtained from a small number of sub-constructs, as further substantiated by Corston and Colman (2003). In order to determine the number and nature of the latent factors responsible for most of the covariance in the current data, the process followed in this study consisted of the following more complex steps which is formulated within the algorithm for data reduction in SPSS (version 17). According to Brown *et al.* (2012):

- Step 1 is to calculate the covariance matrix from the raw data;
- Step 2 then computes a correlation matrix by transforming the aforementioned covariance matrix (in this case, referred to as matrix, A);
- Step 3 is then to extract eigenvectors (in this case, x_i) from the aforementioned correlation matrix by the method of successive squaring;
- Step 4 obtains eigenvalues (λ_i) from each eigenvector (x_i) computed, based on the characteristic equation, $A.x_i = x_i. \lambda_i$ (the equation shows that there exists a unique eigenvalue in each linear transformation);
- Step 5 finally generates the factor loadings (f_i), by a normalisation of the eigenvectors (x_i) to λ_i, that is, $f_i = x_i (\sqrt{} \lambda_i)$;
- Step 6 finalises the results of the above computations, and are displayed in terms of the factor product matrices and factor residuals.
- Step 7 is required to examine the diagonal of the final matrix of residuals to determine the amount of variance that was not accounted for in each variable, and also to examine the off-diagonal elements so as to determine how well the level of correlations amongst each variable pair was subsequently reproduced by the factor solution; and
- Step 8 creates the final matrix of factor loadings (F) from the factors (f_i) extracted. The factor loadings were used as the basis for imputing labels to the explanatory dimensions of the construct under investigation.

The technique therefore uses linear combinations of the empirically obtained variables to explain the sets of observations in the dataset.

5.2.1 Sample size rationale used for the factor analysis

The initial challenge in planning a factor analysis is to determine the size of the sample frame that is likely to result in the most stable solution. Unfortunately, the literature contains a multitude of divergent opinions on the question of the ideal sample size required for factor analysis, which made the selection of an appropriate technique difficult (Gorsuch, 1983; Hayton *et al*., 2004; MacCallum *et al*., 1999). Nonetheless, the main problem to be addressed in this study was in fact, understanding the impact that sampling error, rather than the sample size *per se*, may have had on the final factor analytic solution. Although sampling error can be directly related to sample size, sampling error can be mitigated even when the sample is relatively small, provided that the sample's elements are of a high quality (Haworth, 1996). A high quality final factor solution was deemed critical for this study, because the accuracy and validity of subsequent analyses required a strong foundation. Thus it was important to understand the impact that the final sample frame would have on the factor structure.

Typically, sampling error influences the estimation of factor loadings because the sample covariance of the measured variables is expressed as a function of the population's common and unique loadings (MacCallum *et al*., 1999). This then implies that prescriptive sample sizes, as a ratio of the population size, or in terms of the number of items, might be misleading. The methodology chapter discussed the various rules of thumb proposed by some authors to facilitate the determination of the ideal sample size (see Section 4.10.1). For instance, according to Cooper and Schindler (2003), a sample five times the number of variables is sufficient to produce relatively stable factor solutions, whereas Comrey and Lee (1992:67) suggest that between 200 and 300 participants in a survey using factor analysis should be regarded as only a "mediocre" sample size. Because of these divergent expert opinions, the various rules of thumb have all had their fair share of critics (MacCallum *et al*., 1999). On the basis of the suggested sample sizes discussed above, the final cohort for this study consisted of 229 participants, which is more than the 210 participants that would be five times the number of item variables – a rule of thumb in many similar studies, as recommended originally by Gorsuch (1983) and later by Cooper and Schindler (2003).

MacCallum *et al*. (1999) argue that if the communalities between the variables are relatively high, a small sample size will produce high quality sample factor solutions (in other words, effective replicability and recovery of population factors). Communalities refer to the proportion of each variable's variance that can be explained by the factors (Ledyard, 1966); hence, high communalities in the current study indicated that the final sample, in terms of the validity of the questionnaire items, was highly stable. Using a similar method, Corston and Colman (2003) found that even analysing only a small portion of Thurstone's (1947) data set did not significantly affect their own results.

The average initial communalities of the 42 items which operationalised the main research construct were computed at 0.630. MacCallum *et al*. (1999) consider communality strengths of above 0.60 to be very good. In summary, it appears from various findings reported in the literature that when communalities are relatively good, and the content of the hypothetical framework has been validated (as is the case in the present study), sampling error is reduced (Corston & Colman, 2003). Therefore, using between 200 and 300 returns can produce high quality sample factor analytic solutions. It is also reasonable to assume that very small differences would have occurred if the study had used a sample of 500 versus a sample of 200. The first step in the research process was to validate the content of the construct by using a panel of experts. This choice may be the reason for strengthened communalities amongst the variables. The item communalities (h^2) are reported in Tables 32 and 33.

5.2.2 Factor analytic computation

Using SPSS Version 17.0, in the various rounds of exploratory factor analyses, the 42 items in respect of the content of the hypothetical construct were inter-correlated and analytically rotated to an oblique simple structure by means of the promax method (factors were permitted to optimally correlate with one another). The rotations were raised to a Kappa power of 4, which Hendrickson and White (1964) consider appropriate for most analyses in the social sciences. Kappa is a statistical parameter, which was used to control the calculation of the promax rotation (Morgan, *et al.*, 2007).

Gorsuch (1983) found that when values of Kappa are high, correlations among factors tend to be high, leading to a simpler structure of the loadings. It was concluded that raising the loadings to a power of 4 produced an optimised solution for the dataset, which resulted in a simpler structure with the lowest correlation amongst the factors. An inter-correlation matrix consisting of 1 764 variables was subsequently produced. However, due to its size, this matrix is not reported here.

Prior to the aforementioned factor analytic choice, two diagnostic tests confirmed that the data passed the minimum criteria required to conduct an initial exploratory factor analysis. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (which assesses whether the partial correlations among variables are small) and Bartlett's test of sphericity (which evaluates the null hypothesis that the variables in the population correlation matrix are uncorrelated) both produced satisfactory results for the current data set.

The calculated KMO value of 0.927 was greater than 0.70, which is the normal cut-off recommended for a factor analysis to proceed (Gorsuch, 1983). Similarly, Bartlett's test of sphericity [χ^2 (861) = 6896.895, $p < 0.001$] was significant and therefore confirmed that the properties of the inter-correlation matrix of the 42 item scores were suitable for factor analysis.

Subsequently, using Kaiser's (1961) criterion, principal axis factoring (PAF) postulated seven factors, where the initial eigenvalues (greater than unity) accounted for 66.248% of the variance in the factor space (see Table 31). The extraction was conducted on a correlation matrix, after conversion from a covariance matrix of the raw data. Therefore, the variables were standardised and the total variance (100%) would be found from the remaining items. The loadings and cross-loadings of the items associated with the seven eigenvalues greater than one were then examined. All items that had loadings of less than 0.40 or which were found to have a high cross loading on more than one factor were deleted. A cut-off point of 0.40 is the criterion generally used in the social sciences (Welman & Kruger, 1999). The loadings used in this round of analysis are not reported, due to space constraints. Items whose properties appeared extremely similar were also discarded, in line with the recommendations of Tabachnick and Fidell (2007). In total, seven items were deleted (see Table 32).

Finally, 35 items in a clean matrix were subjected to a second round of exploratory factor analysis with promax rotation and Kaiser Normalization (which is generally the default during rotation). This method decreased the standard errors of the loadings for variables with small communalities.

Table 31: Variance explained by eigenvalues greater than one (42 items)

Initial eigenvalues				Extraction sums of squared loadings	
n	Total	% of variance	Cumulative %	Total	% of variance
1	16.431	39.122	39.122	16.431	39.122
2	3.872	9.219	48.341	3.872	9.219
3	2.174	5.177	53.518	2.174	5.177
4	1.888	4.495	58.013	1.888	4.495
5	1.226	2.919	60.932	1.226	2.919
6	1.174	2.795	63.726	1.174	2.795
7	1.059	2.522	66.248	1.059	2.522

In general, most statistical software programs, and more specifically, the SPSS factor analysis program, defaults to Kaiser's (1961) criterion (eigenvalues greater than unity) for factor extraction. This method and other factor retention methods were examined further. Substantive evidence was found in the literature that demonstrated that surprisingly, Kaiser's criterion was correct only 22% of the time (Hayton *et al*., 2004), which makes it problematic as a factor retention method for scale development at this level.

Hayton *et al.*'s (2004) data also show that Kaiser's criterion will only become accurate when the sample size approaches infinity, and therefore it was expected that any sampling error in the current study would generally produce more factors in the sample space. Deleting some items and subjecting the remaining variables to further rounds of factor analysis assisted in reducing the problems associated with factor over-estimation.

Table 32: Statements deleted in the first round of exploratory factor analysis

Item	Statement
Q35	The standard operating procedures (SOPs) for learning to fly this aircraft are adequate.
Q37	The simulators my company trains its pilots in are in good condition.
Q43	I learn better when I work as a member of the crew.
Q46	I tend to communicate well with my simulator partner.
Q47	The instructor is committed.
Q48	Instructors are very similar in how they teach pilots to fly this aircraft.
Q59	I reflect on my learning experience after a simulator session.

5.2.3 Results of the factor retention method

Cattell's scree plot and Kaiser's (1961) criterion were used to determine tentatively the number of factors to retain, as these appeared to be the most common methods reported in the literature. In Figure 24, the graphs plot the eigenvalues against each component number. From the third or fifth factor onwards, it appears that the plot has flattened, which suggests that each successive factor accounts for less and less of the total variance explained. According to Cattell (1966), factors before the one starting the bend or elbow where the plot levels off should be retained. Inspection of the curve indicating the number of factors to retain showed that the graph was very difficult to decipher, due to the ambiguity in the shape of the curve. Unfortunately, the scree test has been shown to suffer from a degree of subjectivity. For instance, Hayton *et al.* (2004) point out that when there are no clear breaks, or, as in this case, when there are two or more apparent breaks, confusion can arise.

In the current study, this confusion made it difficult to determine unambiguously the optimum number of factors to keep which can accurately describe the latent structure of the measured construct. A further problem was that Kaiser's method appeared to overestimate the number of factors to retain, because six eigenvalues exceeded unity for 35 items (see Table 33).

Figure 24: Thirty-five item scree plot

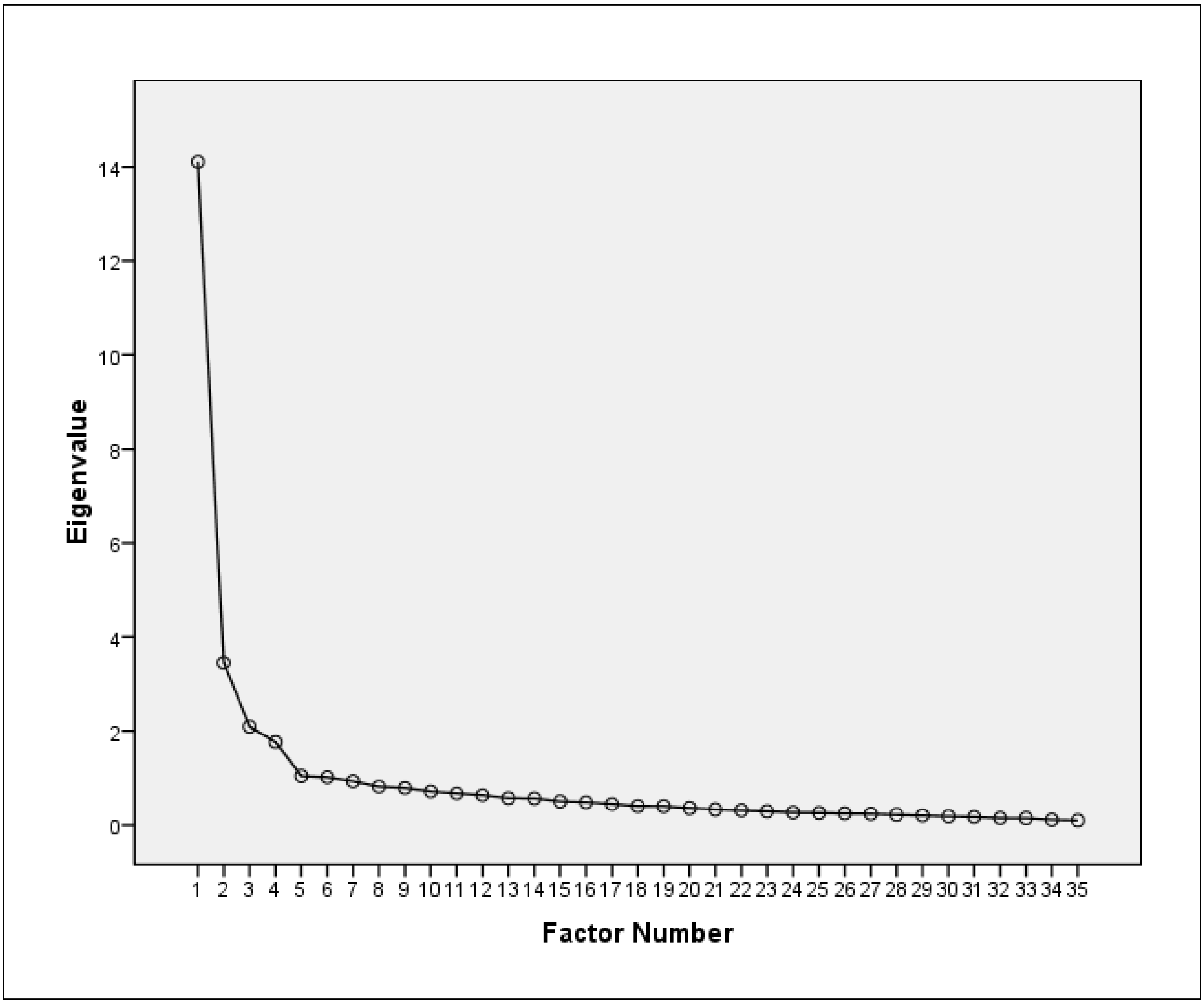


Table 33: Variance explained by eigenvalues greater than one (35 items)

	Initial eigenvalues			Extraction sums of squared loadings		
VAR	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	14.105	40.301	40.301	13.713	39.180	39.180
2	3.454	9.869	50.170	3.148	8.994	48.175
3	2.092	5.978	56.148	1.727	4.934	53.109
4	1.767	5.048	61.196	1.454	4.156	57.264
5	1.044	2.984	64.180	0.639	1.825	59.089
6	1.018	2.909	67.089	0.611	1.745	60.834

It was observed that the first attempt at exploratory factor analysis overestimated the factors in the real test space for this data set using traditional methods. One source of this overestimation may be possible sampling error or differentially skew items (Hendrickson & White, 1964). According to Hayton *et al*. (2004), some factors may have eigenvalues greater than one only because they originate from a finite sample, as opposed to an infinite population. In an infinite population, initial eigenvalues tend to be greater than one, whilst later eigenvalues tend to be smaller than one. Possibly, the most important decision a researcher developing a scale and using factor analysis has to make, is deciding on the number of factors to retain (Glorfeld, 1995; Watkins, 2006).

Accurate factor retention was considered to be a core requirement for the overall success of this study. Both the scree plot and Kaiser's method were inconclusive for the current dataset. To determine the maximum number of significant factors, which would reasonably explain the variability of the main research construct, a modified version of Horn's (1965) parallel analysis (PA) based on a Monte Carlo simulation, was conducted, as recommended by Velicer (1976).

Hayton *et al*. (2004:194) explain that the "rationale underlying PA is that nontrivial components from real data with a valid underlying factor structure should have larger eigenvalues than parallel components derived from random data having the same sample size and number of variables". A comparative method for exploration in factor analysis is considered one of the most accurate factor retention methods available (Rencher, 2002). Unfortunately, most statistical software packages do not offer Horn's (1965) procedure to determine the maximum number of factors to retain (Watkins, 2006). For the current study, the computational algorithm created by Brian O'Connor (2000) was used; and by manipulating this complex syntax code within the SPSS editor the required analysis was achieved.

Figure 25: O'Connor plot of the actual, mean and permutated eigenvalues

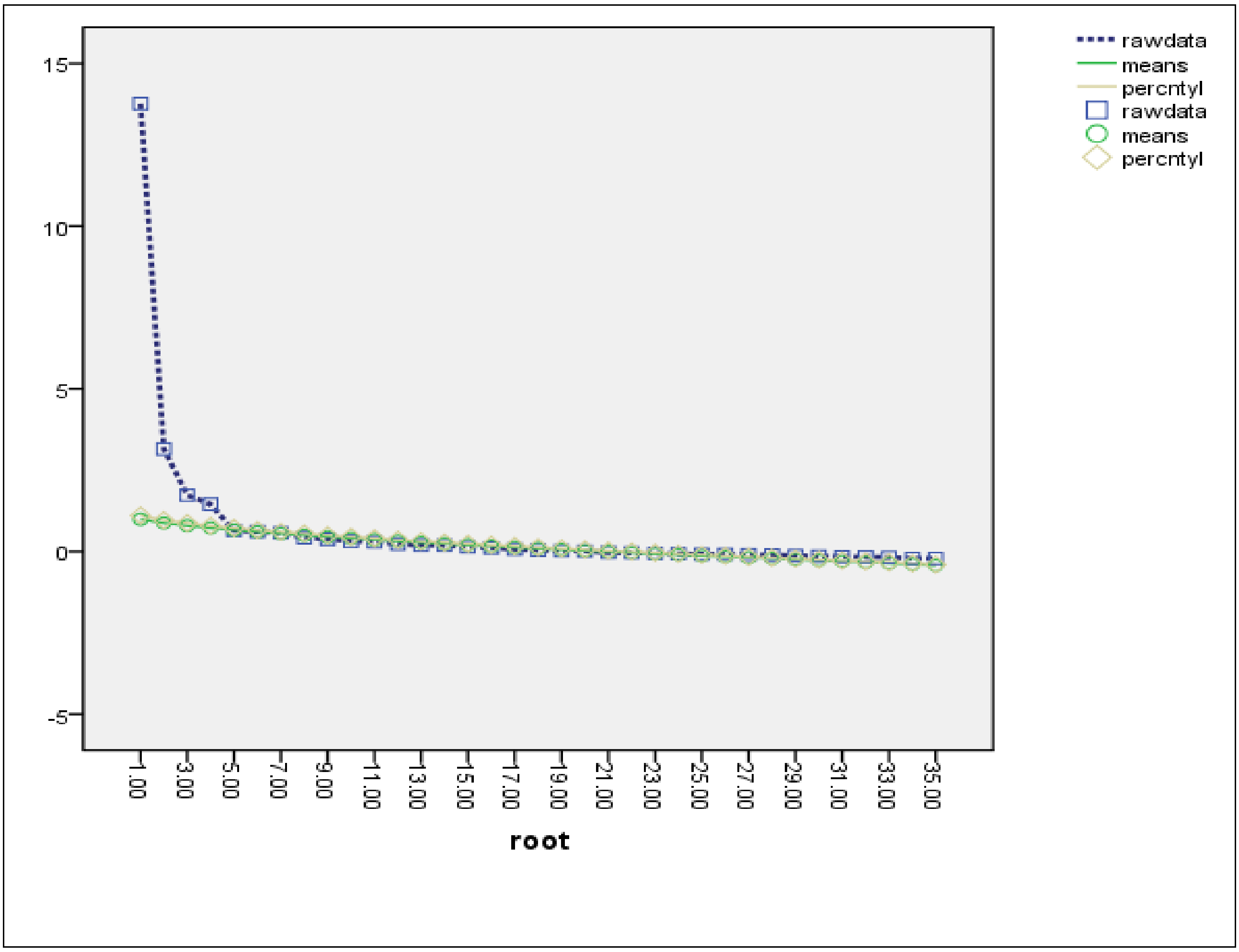


Table 34 and Figure 25 show the actual eigenvalues drawn from the real data space of a 35-item AATC-Q for 229 participants, together with the mean and 95th percentile eigenvalues extracted from 1 000 permutations of the original data set generated in a Monte Carlo simulation. In this simulation, in cases where the raw data are not normally distributed or where they do not meet the assumption of multivariate normality, permutations are deemed more accurate and relevant than randomised data (O'Connor, 2000)

Table 34: Actual and permutated eigenvalues on 35 items based on O'Connor's (2000) algorithm

Root	Actual eigenvalue	Mean eigenvalue	95th percentile eigenvalue
1.	13.763320	0.991263	1.110948
2.	3.152418	0.877747	0.964691
3.	1.734102	0.794188	0.876276
4.	1.451794	0.723979	0.792523
5.	0.649697	0.660416	0.722079
6.	0.612930	0.601232	0.659002
7.	0.574928	0.548531	0.604958
8.	0.422722	0.498339	0.553514
9.	0.381465	0.450765	0.499703
10.	0.320640	0.404774	0.451020
11.	0.297724	0.360918	0.404150
12.	0.233578	0.318256	0.361014
13.	0.199624	0.278398	0.317899
14	0.196496	0.239460	0.277779
15.	0.140422	0.202216	0.237528
16.	0.116488	0.166291	0.204303
17.	0.074143	0.129561	0.166742
18.	0.042527	0.094801	0.129188
19.	0.024271	0.060795	0.095014
20.	0.014329	0.027945	0.060390
21.	-0.023788	-0.004203	0.027235
22.	-0.035305	-0.035648	-0.003258
23.	-0.053125	-0.067442	-0.036161
24.	-0.065674	-0.097213	-0.068280
25.	-0.078143	-0.127373	-0.099662
26.	-0.084013	-0.157061	-0.130434
27.	-0.096836	-0.185802	-0.160826
28.	-0.105393	-0.213957	-0.188567
29.	-0.124817	-0.242558	-0.218763
30.	-0.144008	-0.270705	-0.247826
31.	-0.159285	-0.299199	-0.276946
32.	-0.164820	-0.328068	-0.304783
33.	-0.175846	-0.357945	-0.333908
34.	-0.209793	-0.390079	-0.363294
35.	-0.214610	-0.428936	-0.399151

The permutation algorithm based on Castellan's (1992) recommendation was used to compute the data set. The algorithm ensures that for *N* elements there are *N!* permutations, and each one is equally likely (Table 34). In accordance with the previous extraction methods, a principal axis or common factor analysis was requested for the parallel analysis. A difference can be clearly observed between the eigenvalues produced by the SPSS programme and those produced using the O'Connor (2000) algorithm. The SPSS statistical software program computes a correlation matrix, which differs slightly from that being used in the syntax of the parallel analysis. For instance, Kendall correlation coefficients are generally used for ordinal variables when assumptions of normality are violated, instead of Spearman correlation coefficients, as in the case of normalised data.

Furthermore, when the matrix is not Gramian, in that squared multiple correlations are inserted into the diagonal of the matrix as communalities; a principal factor extraction tends to produce a number of negative eigenvalues (Comrey & Lee, 1992). According to Gentle (2007:288), a Gramian matrix is "a (real) matrix A such that for some (real) matrix B, $A = B^TB$". Because B^TB is symmetric, any non-negative definite matrix is Gramian. In addition, "each element of a Gramian matrix is the dot products of the columns of the constituent matrix" (Gentle, 2007:228). It was determined that in the present SPSS algorithm, the issue is automatically resolved by adjusting the off-diagonal elements of the correlation matrix before factorisation, which is not the method adopted in O'Connor's (2009) program. However, making an adjustment in the off-diagonal elements will, in turn, result in an underestimation of the values of communalities. The parallel analysis syntax uses a more accurate shuffling transformation technique, as recommended by Castellan (1992), for the generation of comparable permutated matrices. According to Castellan (1992:72), the algorithm, "initially generates a random integer between 1 and N, swaps the first element with the generated element, then generates a second random integer between 2 and N, swaps the second element with the second generated element, and so forth". Coding in this manner in order to generate permutated data sets is regarded as efficient, in that it requires only N-1 random integers to permutate the array of N elements; furthermore, the elements are sampled without replacement (Castellan, 1992). In the current study, eigenvalues after the 21st variable in the O'Connor output were negative (dropped below zero). This result is due to the difference between the computation formulae adopted by the SPSS programmers for factor analysis and that used in the O'Connor syntax. Both formulae are acceptable for the purposes of exploratory factor analysis (Rencher, 2002). In addition, an iterative principal axis factoring technique (as in a Monte Carlo type simulation) is more likely to produce negative eigenvalues than a standard principal components method (Watkins, 2006). Also, anomalies can occur when eigenvalues and eigenvectors are computed after first estimating the initial values of the covariance matrix, as in the case of a principal axis factoring method (Rencher, 2002). The results set out in Figure 25 and Table 34 suggest that there are, at most, four eigenvalues greater than those generated by the parallel analysis for both the mean and the 95th percentile criterion. The raw data eigenvalues greater than chance ranged from 13.7 to 1.4. The 95th percentile eigenvalues for this range fell between 1.1 and 0.79. Therefore, four factors were deemed statistically significant ($p < 0.05$), suggesting that a maximum of four components is possible in the factor space of the current data set (Velicer, 1976), and not seven, as was originally postulated using the Kaiser (1961) criterion. O'Connor (2000) suggests that the eigenvalues from a parallel analysis should be used to determine the maximum real data eigenvalues that can occur beyond chance, and additional procedures should then be used to trim trivial factors. In other words, the current parallel analysis indicated the maximum possible number of factors, but it did not impose the final number of factors to retain, as Hayton *et al.* (2004) also found.

5.2.4 Finalised factor analytic solution

The aim in the final factor solution was to obtain a simple structure, after an exhaustive retention process (see also Appendix C, for the final 3-scale solution). Only by determining the probable number of real factors that could exist in the variable space from a Monte Carlo simulation could a maximum four-factor solution be requested. Unfortunately, a four-factor structure resulted in generalised Heywood cases, which, according to Harman and Yoichiro (1966:563), are mathematical anomalies that occur "when a correlation matrix yields a suitable factor solution with several common factors for which one of the communalities exceeds unity". Such cases indicate that the solution may be unstable, possibly because are too many or too few factors. Therefore, this result suggested that the solution was far from simple at this point. With this in mind, a three-factor solution was then requested, because "if the researcher is interested in using only demonstrably reliable factors, the fewest possible factors are retained” (Tabachnick & Fidell, 2007:646).

Table 35: The factor loadings and communalities (h^2) for the principal factors extraction and promax rotation for the final 33-item cohort

Item	Statement	Factor			
		1	2	3	h^2
Q29	Training on this aircraft is well organised.	0.809	-0.041	-0.021	0.708
Q27	Training on this aircraft is professional.	0.808	0.082	0.015	0.767
Q23	My company’s training produces world class pilots.	0.785	0.121	-0.078	0.770
Q24	Training at my airline is in line with company goals.	0.785	0.115	-0.057	0.725
Q38	The airline is very supportive of its pilots’ learning requirements for this aircraft.	0.763	-0.059	0.108	0.703
Q34	There is sufficient training guidance from the company.	0.762	-0.143	0.109	0.667
Q28	Management follows the rules and regulations appropriately.	0.755	-0.053	-0.102	0.646
Q39	My company’s culture supports training for new technology aircraft.	0.731	-0.008	-0.045	0.594
Q30	I understand what the company expects of me when training.	0.700	0.198	0.000	0.706
Q26	My company has talented people in training.	0.695	0.012	0.060	0.633
Q33	If I had to experience a problem in training, it’s easy for me to appeal.	0.686	-0.250	0.291	0.681
Q25	I know what my company’s training goals are.	0.678	-0.036	0.146	0.697
Q31	Training at my airline produces safe pilots.	0.644	0.217	-0.043	0.670
Q40	There is sufficient feedback about my training on this aircraft.	0.606	0.029	0.090	0.546
Q42	My company uses only current training material.	0.589	-0.106	0.035	0.428
Q41	Training is in line with civil aviation regulations.	0.586	0.281	-0.078	0.595
Q32	The airline gives its pilots an appropriate amount of preparation work for training.	0.531	-0.028	0.189	0.525
Q45	My instructor is willing to listen.	0.528	-0.028	0.277	0.639
Q50	Pilots are in direct control of the training outcome.	0.522	-0.111	0.397	0.696
Q36	I’m given sufficient time to prepare for training on this aircraft.	0.499	0.046	0.192	0.558
Q61	It’s a good idea to know more than what is required.	-0.341	0.759	0.340	0.689
Q52	I try never to be late for a training session.	0.208	0.746	-0.257	0.828
Q53	I co-operate when training in a simulator.	0.209	0.745	-0.238	0.810
Q62	I aim to gain a deeper understanding of this aircraft.	-0.312	0.734	0.328	0.773
Q51	Preparation improves performance.	0.254	0.655	-0.065	0.733
Q60	I read to understand so as to gain a deeper understanding of this aircraft’s systems.	-0.218	0.626	0.258	0.672
Q55	I have a positive relationship with my colleagues.	0.224	0.586	-0.011	0.619
Q44	I operate well as a crew member in the simulator.	0.182	0.521	0.148	0.595
Q58	I enjoy studying the technical aspects of the aircraft.	-0.089	0.493	0.316	0.594
Q63	I’m comfortable undergoing training for this aircraft.	0.182	0.130	0.594	0.601
Q57	I’m in control of the outcome of a training session.	0.326	-0.017	0.529	0.656
Q64	I can control my anxiety so as to perform well in training.	-0.073	0.103	0.523	0.497
Q49	The instructors on this aircraft don’t overload us with information.	0.398	-0.076	0.461	0.591
Eigenvalues = 14.105; 3.454; 2.092; % variance = 40.301; 4.869; 5.978; Cumulative % = 56.148					

A quantitatively superior solution was found by retaining only three factors (with no further Heywood anomalies). However, in the final 35-item cohort, Item 54 (“After training I feel a sense of mastery”) and Item 56 (“Pilots who come prepared have no problems training for this aircraft”) had weak loadings (< 0.40) in the final rotated matrix. They were therefore also discarded from the final factor solution.

The factor loadings and the communalities (h^2) of the 33 items, as well as the total variance explained by the different factors, are depicted in Table 35. The factor loadings are set out in descending order for easier interpretation. Inspection of these results suggests that the three factors were well determined and that the three factors explained approximately 56.148% of the total variance in the data.

According to the results, the factor scores of the factor solution ranged from excellent to fair, with factor scores varying from 0.809 to 0.461:

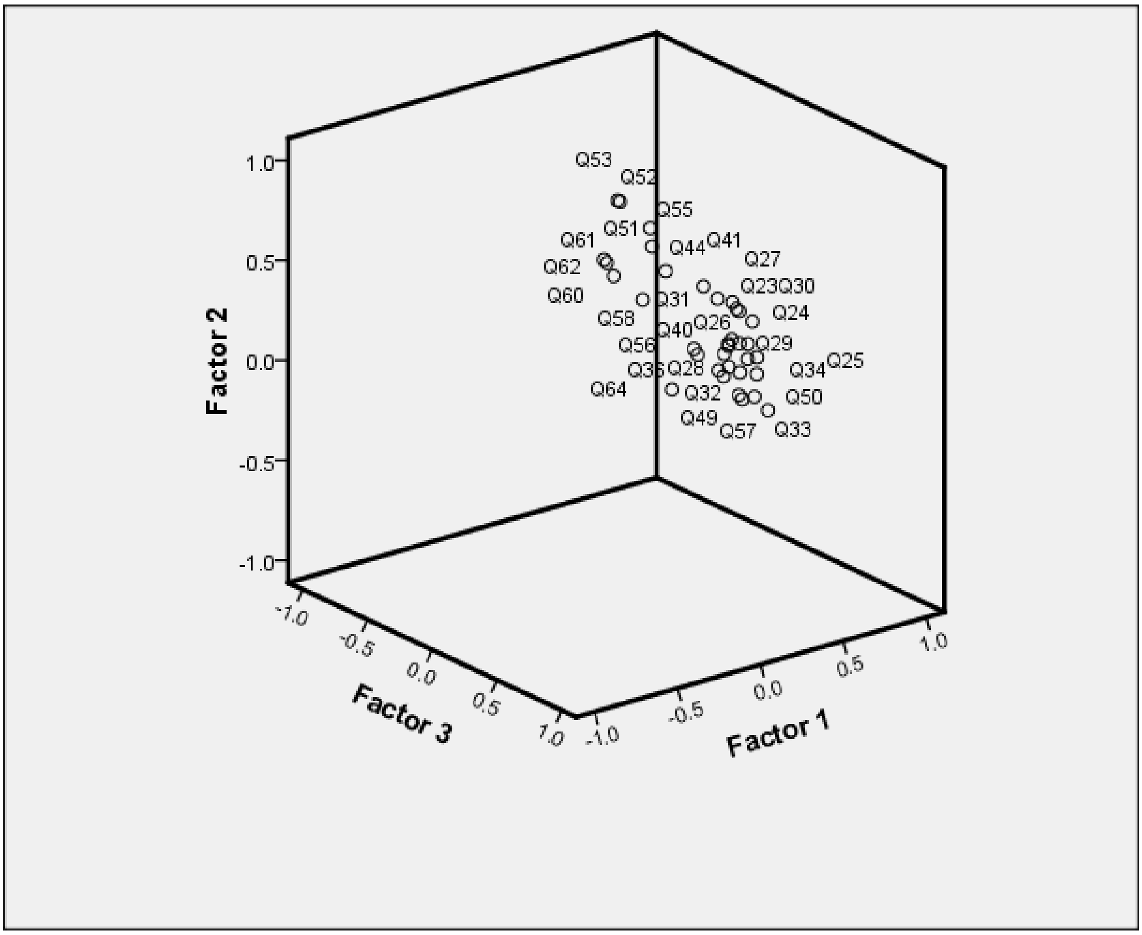
- 0.809 to 0.499 for Factor 1;
- 0.759 to 0.493 for Factor 2; and
- 0.594 to 0.461 for Factor 3.

Comrey and Lee (1992:203) suggest that loadings in excess of 0.70 be considered “excellent”, 0.63 or more “very good”, 0.55 or more “good”, 0.45 or more “fair” and 0.32 or more “poor”. Loadings lower than 0.32 should be disregarded. High factor loadings are a clear indication that a variable is a “pure” measure of the factor (Tabachnick & Fidell, 2007:649).

In the current study, the average communality loadings (h^2) of 0.701 to 0.586 indicated that the items define the factors relatively well, representing high levels of consistency (Leech *et al.*, 2005). In this kind of context, Tabachnick and Fidell (2007:660) point out that “[c]ommunalities indicate the percentage of variance in a variable that overlaps variance in the factors”. Comrey and Lee (1992:12) define the communality for a variable in factor analysis as “the sum of squares of the factor loadings over all the factors”. The final factor loadings and communalities that emerged from the analysis were dependent on the original estimated communality values used in the early correlation matrix (Comrey & Lee, 1992). Therefore, high communalities (see Table 35) indicate homogeneity and the factorial purity of the final scale (Appendix C).

The grouping of variables in the rotated space also indicates that the factors were sufficiently described by their items, because a “clustering of variable points reveals how clearly defined a factor is” (Tabachnick & Fidell, 2007:647). The linear combinations of the variables are presented in the rotated factor pattern matrix (see Figure 26). However, only one clear factor cluster (Factor 1) can be seen within the rotated space. The remaining two factors may not be as clear, because values of the loadings are based on the common variance, which is generally lower than the total variance.

Figure 26: Factor plot in the rotated space



In addition, the internal consistency of the factor solution was verified by calculating the Squared Multiple Correlations (SMCs). This index indicates that the variable data points were fixed in the factorial space to a high degree of certainty (Bentler & Raykov, 2000), as set out in Table 36.

The importance of each factor was assessed by the percentage of variance that it represented. The Sum of Squared Loadings (SSLs) from the rotated factors is the redistributed variance during rotation. Each of the factors accounted for between 6% and 40% of the covariance. As expected, the first factor (F1) accounted for the bulk of the covariance.

The SMCs values and the intercorrelation between the three factors are depicted in Table 36. The results suggest that the three factors intercorrelated significantly with one another (r = 0.242 to 0.446). The strength of the correlations indicates that the three factors are closely related in measuring the constructs associated with the advanced automated aircraft training climate. Although the relatively high intercorrelations may also suggest overlapping variability, the SMCs nonetheless indicate that all the factors were sufficiently defined by the relevant items.

Table 36: Item regression model and factor correlations

SMC Model (Factor)	R	R Square	Adjusted R Square	Standard error of the estimate
1	0.995	0.990	0.989	0.10083248
2	0.991	0.983	0.982	0.12711818
3	0.900	0.809	0.806	0.40278467
	Factor correlations			
Factor	F1		F2	F3
1	1.000		0.446	0.431
2	0.446		1.000	0.242
3	0.431		0.242	1.000

The SMCs or R^2 values in Table 36 were calculated by means of the regression method described by Tabachnick and Fidell (2007). Items were taken as predictors of the factor constructs to produce the SMC models shown. The high R^2 values, ranging from 0.809 to 0.990, between the item scores and the factor scores testify to a good fit between the item scores and the latent factors. The R^2 values also represent the proportion of shared variance between each of the items and its related factor. The results indicate that the items could account for between 80% and 98% of the variation in the three factors. The model provides the necessary evidence to demonstrate that the items are significant explanatory variables of the latent constructs of pilots' perceptions of the advanced aircraft training climate.

The results of the analysis discussed above suggest that the items of the questionnaire exceeded the required adequacy in measuring the factors that they were related to. The item clusters were used as a guide for the factor labelling process.

5.3 SCALE LABELLING AND FACTOR DESCRIPTION

In order to communicate the nature of the underlying constructs, factors must be appropriately labelled (Ledyard, 1966). From an inspection of the content of the sub-scales, it is possible to gain a deeper understanding of the nature of the underlying factor constructs (Comrey & Lee, 1992). In addition, factors are represented by different sets of variables and therefore each factor should be well differentiated from the others.

“The variables that have high loadings on the rotated factor are studied carefully to determine what they share in common” (Comrey & Lee, 1992:11). The description and naming of the latent factors that accounted for most of the variability in the main research construct was based on the three to five statements with the highest explanatory connotations within each grouping such that:

- *Factor 1*

  This factor essentially relates to the organisational aspects of the training climate. Items from both the macro domain (the airline) and the intermediate (instructor-trainee) domain loaded substantively onto Factor 1. Essentially, the factor expresses a component of the theoretical construct in terms of the efficiency, effectiveness and professionalism of both the company and its flight instructors. The elements dominating this factor relate to organisational co-ordination, trainee support, rules, regulations, sufficient learner feedback and guidance. This factor is referred to as *Organisational Professionalism*.

- *Factor 2*

  This factor contains elements representing the micro level of analysis (the person). The factor predominantly reflects the individual trainee's ability and eagerness to learn and understand complex concepts relating to advanced aircraft. Learning the aspects of a complex technology is regarded as a structured and iterative quantitative increase in knowledge. The fundamental aspects of this factor relate to an individual's learning approach, preparedness, and willingness to participate and co-operate in training to gain a knowledgeable and workable understanding of the advanced aircraft. This factor has been labelled *Intrinsic Motivation*.
- *Factor 3*

  The third factor represents an individual trainee's own perceived level of control in terms of stress levels and decision making, regarding the training required to operate an advanced automated aircraft. Four items found in the micro domain (the person) loaded meaningfully onto this factor. The principal elements of Factor 3 relate to the levels of perceived comfort experienced by trainees during training, their belief in their ability to control the outcome of a training session, their capacity to control their levels of stress (eustress or anxiety) in order to perform well, and ultimately their grasp of the amount of information required to cope with their training (intelligent decision making). This factor is referred to as *Individual Control of Training Outcomes.*

Comrey and Lee (1992:11) suggest that a researcher go beyond simply doing a factor analysis and labelling the factors, because a factor analysis should be considered "a way of generating hypotheses about nature". The derived factor constructs and rotated factor matrix can therefore be considered as only one interpretation of the latent behaviour of the present population (Appendix E).

5.4 RELIABILITY ANALYSIS

“The reliability of a measuring instrument is defined as its ability to consistently measure the phenomenon it is designed to measure” (Ho, 2006:239). However, it is expected that data variables will not be perfectly reliable because they are not perfectly correlated (Comrey & Lee, 1992). In order to determine the level of internal consistency (the interrelatedness of a set of items and the extent to which the items in a scale measure the same construct), Cronbach’s coefficient alpha was computed for each of the sub-scales. In addition, the mean inter-item correlations were computed to assess the internal homogeneity and unidimensionality of the item clusters in each factor scale, as recommended by Pett *et al*. (2003). Furthermore, the mean, standard deviation, skewness, and kurtosis, for each of the items were calculated to assess the distribution of the reposes of the present sample on the 33 item AATC-Q

As described in Tables 37 to 39, the Cronbach’s coefficient alphas (α) of all three factors were relatively high and exceeded the recommended threshold of 0.70, which is the accepted standard requirement for levels of internal consistency in social sciences studies of this nature and for the number of items in a scale (Comrey & Lee, 1992; Cortina, 1993; Cronbach, 1951; Morgan & Griego, 1998; Nunnally & Bernstein, 1994). Furthermore, the findings indicate that each item contributed significantly to the high reliability coefficient within each factor. It was also apparent that none of the items adversely reduced the value of alpha if the item was removed from the cluster.

The high Cronbach coefficient alphas were also interpreted as an indication that there is very little variance specific to individual items. In other words, the sets of items were found to conform to Cronbach’s (1951) original definition of *equivalence*, where the value of Cronbach’s alpha takes into account all the information contained in the items, namely the number of items, their variance and covariance.

Table 37: Reliability and item statistics for Factor 1: *Organisational Professionalism* (n =229)

Item statistics	Sk	Ku	Mean	Standard deviation	Corrected item total correlation	Cronbach's alpha if the item is deleted
Q23	-1.42	2.17	5.97	1.162	0.776	0.948
Q24	-1.22	2.26	6.04	1.023	0.781	0.949
Q25	-1.18	1.23	5.67	1.355	0.722	0.949
Q26	-1.60	3.65	6.03	1.151	0.700	0.949
Q27	-1.81	4.19	6.15	1.066	0.828	0.948
Q28	-1.29	1.23	5.46	1.497	0.668	0.950
Q29	-1.42	1.99	5.65	1.370	0.767	0.948
Q30	-1.43	2.33	6.08	1.079	0.759	0.949
Q31	-1.41	2.17	6.00	1.116	0.693	0.949
Q32	-1.21	1.14	5.56	1.449	0.600	0.951
Q33	-1.09	0.69	5.39	1.609	0.701	0.950
Q34	-1.36	1.78	5.65	1.367	0.736	0.949
Q36	-1.61	2.55	5.91	1.383	0.595	0.951
Q38	-1.33	1.53	5.66	1.417	0.783	0.948
Q39	-1.38	1.48	5.75	1.385	0.689	0.949
Q40	-1.25	1.31	5.65	1.288	0.659	0.950
Q41	-2.24	6.07	6.38	1.018	0.649	0.950
Q42	-1.51	2.71	5.90	1.201	0.553	0.951
Q45	-1.53	2.78	5.77	1.275	0.623	0.950
Q50	-0.94	0.57	5.21	1.401	0.637	0.950

Reliability statistics	N of items	Mean inter-item correlation	Cronbach's alpha
	20	0.663	0.952
Scale statistics	Mean	Variance	Standard deviation
	115.86	348.249	18.661

Table 38: Reliability and item statistics for Factor 2: *Intrinsic Motivation* (n = 229)

Item statistics	Sk	Ku	Mean	Standard deviation	Corrected item total correlation	Cronbach's alpha if the item is deleted
Q44	-1.01	0.92	6.16	0.844	0.599	0.865
Q51	-1.89	3.27	6.52	0.820	0.645	0.861
Q52	-3.29	11.20	6.76	0.694	0.624	0.865
Q53	-2.69	6.98	6.74	0.643	0.635	0.865
Q55	-1.17	1.13	6.41	0.724	0.603	0.866
Q58	-1.28	2.39	5.72	1.131	0.560	0.873
Q60	-1.25	2.09	5.92	1.099	0.632	0.864
Q61	-1.73	3.83	6.31	0.944	0.678	0.858
Q62	-1.27	2.15	6.17	0.926	0.711	0.855
Reliability statistics				N of items	Mean inter-item correlation	Cronbach's alpha
				9	0.632	0.877
Scale statistics				Mean	Variance	Standard deviation
				56.71	31.910	5.649

Table 39: Reliability and item statistics for Factor 3: *Individual Control of Training Outcomes* (n =229)

Item statistics	Sk	Ku	Mean	Standard deviation	Corrected item total correlation	Cronbach's alpha if the item is deleted
Q49	-1.22	1.38	5.36	1.387	0.544	0.701
Q57	-1.15	1.56	5.52	1.255	0.546	0.692
Q63	-2.23	7.48	6.26	0.968	0.691	0.634
Q64	-1.54	3.08	5.83	1.139	0.452	0.741
Reliability statistics				N of items	Mean inter-item correlation	Cronbach's alpha
				4	0.558	0.750
Scale statistics				Mean	Variance	Standard deviation
				22.97	13.109	3.621

An examination of the reliability analyses in Tables 37 to 39 showed that the corrected item correlations of all three factors indicate that the specific items would make good components of a summated rating scale. The mean inter-item correlations scores on the three factors also satisfy the requirements of homogeneity and unidimensionality suggested by Clark and Watson (1995). An examination of the scores of the mean inter-item correlations in Tables 37 to 39 indicates that the items measured a narrow or well-defined construct (Zeller & Carmines, 1980). The average inter-item correlations for the three factors all yielded exceptionally high values (Factor 1 = 0.663; Factor 2 = 0.632 and Factor 3 = 0.558).

Clark and Watson (1995) suggest that the specificity of the target construct is prominent when the average inter-item correlation exceeds 0.50. Using a panel of experts to initially scrutinise construct items (from an application of the Lawshe method) considerably improved the quality of the items and, more importantly, trimmed invalid items, thereby retaining superior final variables for scale development.

Based on the results reported above, all the items of the three factors in each table were retained as separate indicators to measure airline pilots' perceptions of the training climate associated with advanced automated aircraft. An inspection of the items in Factor 1 (Table 37), Factor 2 (Table 38) and Factor 3 (Table 39) reveals that all the item means are between 5 and 7, with an approximate standard deviation of 0.9 to 1.1, and that all the skewness coefficients are negative, ranging from -0.940 to -3.29. A normal distribution has a skewness index of zero. Thus, the distribution of each factor would have a long "tail" to the left and can be said to deviate somewhat from normality (Morgan & Griego, 1998:49).

Kurtosis determines whether the peak of the distribution is higher or lower than the ideal normal curve. In a normal distribution, kurtosis is equal to zero. An examination of the results revealed high kurtosis values, ranging from 0.57 to 11.20. Here, the positive kurtosis index implies very peaked curves. These deviations from the normal distribution were expected due to the homogeneous nature of the sample. The normality of the data and the implications of non-normal distributions for the selection of statistical procedures are explored in Section 5.6.

5.5 ITEM DISCRIMINATION ANALYSIS

The goal of the study from here on is to understand the nature and finer characteristics of the scaled construct. According to Brown (1983), researchers should ensure that their measurements are as valid and as reliable as possible when phenomena are measured on the basis of psychological scales. An item discrimination analysis was deemed necessary in order to fulfil Brown's (1983) requirement for increased proof of validity and reliability.

A scale is considered more effective when its items differentiate successfully between the proportions of high and low scorers (Cortina, 1993; Leech *et al*., 2005). A well-formulated item should have the predictive power to place top scoring participants into an upper group and lower scoring participants into a lower group.

To determine the level of item-discrimination in the scale of the current study, a discriminant function analysis was used to classify a dichotomous dependent variable based on information from continuous or ordinal independent variables (in this case, the factor items). The latent construct factors for this process were considered independent for the purposes of the analysis (in other words, reversal). Cases were allocated a place in a dichotomous group labelled *Discriminant Category*. The formulated *Discriminant Category* group became the dependent variable for the purposes of this computation. The group contained two dummy variables, 0 and 1. Each placement was based on whether the mean inter-item score was in the upper bound (>5.0, Dummy Variable 1) or the lower bound (<4.99, Dummy Variable 0). Prior to conducting the analysis, however, a matrix scatter plot (see Figure 27) was used to analyse the basic assumptions required for proper model specification.

Figure 27 depicts a comparison of the scatter plots for the upper bound (Dummy Variable 1) and the lower bound (Dummy Variable 0) groups. It appears that the scatter plots for the same variables are similar with regard to their variability for the two groups. Therefore, the assumption of the homogeneity of variance-covariance matrices was met and additional quantitative analyses based on the discriminant function could proceed. Table 40 sets out the results of the discriminant significance tests based on Wilks's lamda for each item. Wilks's lamda is the statistic of choice to separate variable classes (Mardia, Kent, & Bibby, 1979).

Figure 27: Matrix scatterplot for the discrimination of classes

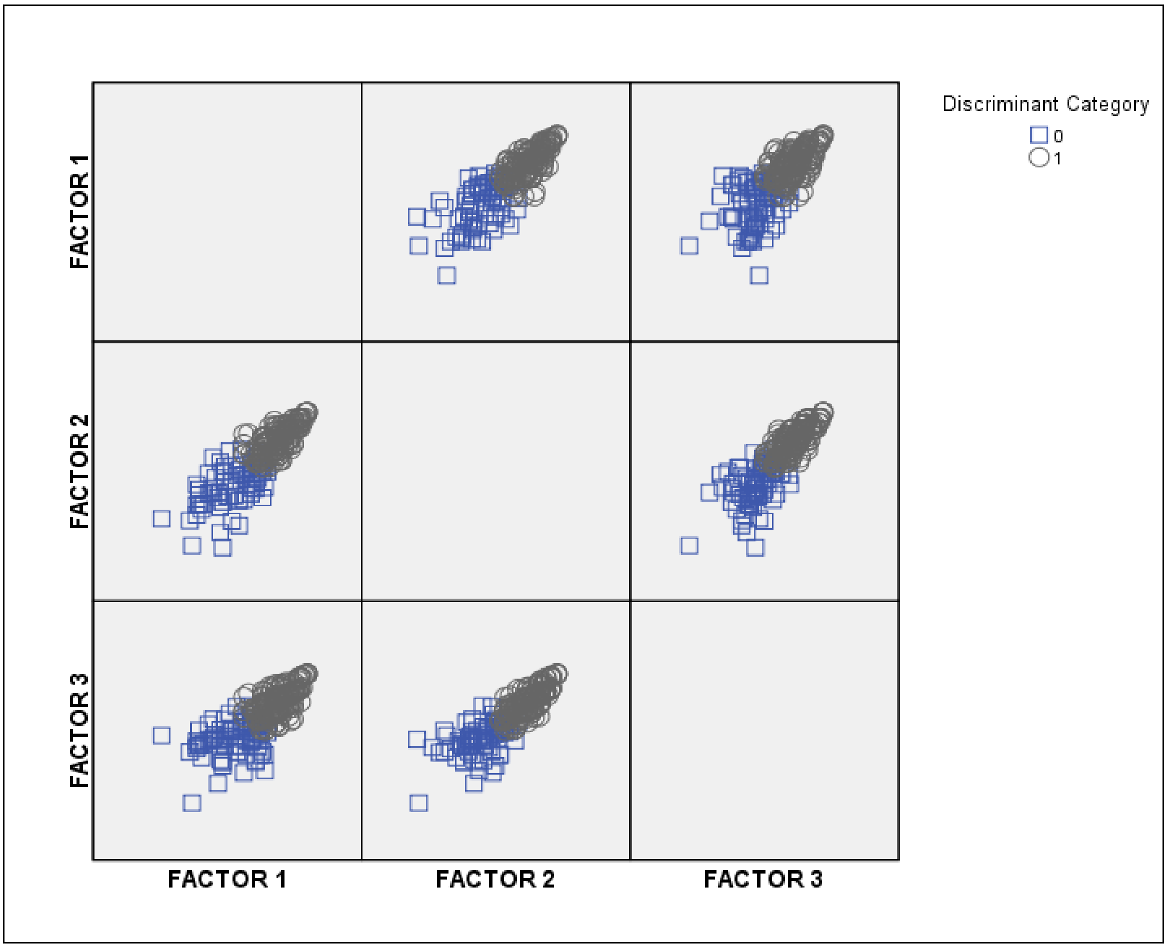


Table 40: Tests of equality of the discriminant group means

Item		Wilks's lambda	F	df1	df2	Sig.
Q23	My company's training produces world-class pilots.	0.820	49.742	1	227	0.000
Q24	Training at my airline is in line with company goals.	0.838	43.956	1	227	0.000
Q25	I know what my company's training goals are.	0.846	41.381	1	227	0.000
Q26	My company has talented people in training.	0.867	34.969	1	227	0.000
Q27	Training on this aircraft is professional.	0.808	54.019	1	227	0.000
Q28	Management follows the rules and regulations appropriately.	0.865	35.433	1	227	0.000
Q29	Training on this aircraft is well organised.	0.841	42.902	1	227	0.000
Q30	I understand what the company expects of me when training.	0.801	56.445	1	227	0.000
Q31	Training at my airline produces safe pilots.	0.827	47.510	1	227	0.000
Q32	The airline gives its pilots an appropriate amount of preparation work for training.	0.911	22.099	1	227	0.000
Q33	If I had to experience a problem in training, it's easy for me to appeal.	0.880	30.986	1	227	0.000
Q34	There is sufficient training guidance from the company.	0.866	35.184	1	227	0.000
Q36	I'm given sufficient time to prepare for training on this aircraft.	0.850	39.908	1	227	0.000
Q38	The airline is very supportive of its pilots' learning requirements for this aircraft.	0.839	43.424	1	227	0.000
Q39	My company's culture supports training for new technology aircraft.	0.879	31.332	1	227	0.000
Q40	There is sufficient feedback about my training on this aircraft.	0.860	37.087	1	227	0.000
Q41	Training is in line with civil aviation regulations.	0.860	37.088	1	227	0.000
Q42	My company uses only current training material.	0.942	13.885	1	227	0.000
Q44	I operate well as a crewmember in the simulator.	0.731	83.422	1	227	0.000
Q45	My instructor is willing to listen.	0.793	59.407	1	227	0.000
Q49	The instructors on this aircraft don't overload us with information.	0.814	51.897	1	227	0.000
Q50	Pilots are in direct control of the training outcome.	0.743	78.615	1	227	0.000
Q51	Preparation improves performance.	0.801	56.292	1	227	0.000
Q52	I try never to be late for a training session.	0.874	32.751	1	227	0.000
Q53	I co-operate when training in a simulator.	0.862	36.387	1	227	0.000

Item		Wilks's lambda	F	df1	df2	Sig.
Q55	I have a positive relationship with my colleagues.	0.779	64.554	1	227	0.000
Q57	I'm in control of the outcome of a training session.	0.719	88.890	1	227	0.000
Q58	I enjoy studying the technical aspects of the aircraft.	0.741	79.141	1	227	0.000
Q60	I read to understand so as to gain a deeper understanding of this aircraft's systems.	0.718	89.171	1	227	0.000
Q61	It's a good idea to know more than what is required.	0.770	67.908	1	227	0.000
Q62	I aim to gain a deeper understanding of this aircraft.	0.700	97.463	1	227	0.000
Q63	I'm comfortable undergoing training for this aircraft.	0.678	107.95	1	227	0.000
Q64	I can control my anxiety so as to perform well in training.	0.711	92.195	1	227	0.000

According to the values set out in Table 40, Wilks's lambda varied from 0.678 to 0.942, indicating that the variables differentiated moderately between the upper and lower groups (Mardia, *et al*., 1979). The F-test of Wilks's lambda shows the variables' contributions that are significant. The F-test significance levels for all the items were good ($p < 0.001$), which in turn indicated that each of the items in the scale was a significant discriminant group predictor by itself.

In addition, the computation of the chi-square test statistic for the data also confirmed the inequality of the location of the mean scores of the upper bound and the lower bound groups (χ^2 [1,127] = 278.347; $p < 0.001$), and attested that the 33 items in the combined scale (predictors) are able to separate the upper and lower discriminant groups effectively and significantly, suggesting that the scale has an appropriate discriminant ability.

5.6 DATA EXPLORATION: ANALYSIS OF DISTRIBUTION

Normality of data is an underlying assumption of parametric statistical testing; therefore testing that the scores are normally distributed and variances of groups are equal is a prerequisite for studies of this nature (Field, 2005). In order to decide on the most appropriate family of statistics to analyse the summated scale scores further, variable score distributions were examined. Table 41 depicts the descriptive statistics of the instrument and some fundamental demographics.

Variables were considered continuous scales when measuring their means, the standard deviations, kurtosis and skewness of scores (because each item was based on a Likert design). The rejection of the assumption of normality can be determined from key statistics. In cases where the skewness and kurtosis are more than 2.5 times the standard error, the assumption of normality of data is violated (Morgan & Griego, 1998:49).

Next, the Kolmogorov-Smirnov and Shapiro-Wilks tests were used to test goodness-of-fit to obtain a statistic (a z-value) and a p-value. Apart from the three scales, the test was also applied to important demographic variables that described the sample. These were further collapsed into categories of the airline company size (small, medium, large) in Table 42 for further scrutiny.

It was important to determine the extent of non-normality amongst the airline carriers in terms of their sizes, because the literature demonstrates that the number of aircraft operated by an airline can directly influence organisational phenomena such as training. For instance, there is a direct correlation between the number of aircraft in a fleet and the preponderance of schedule frequencies, the number of pilots and the resources available (Wilson & Weston, 1989). These results are reported in Table 41.

The results show that the null hypothesis that either the demographic or the factor space samples come from a one-dimensional normal probability distribution should be rejected ($p < 0.05$) for all three factors and for the majority of the demographic variables. The two components of normality (Sk and Ku) were assessed (see Table 41). The results show that the mean was not in the centre of the distribution for either the continuous independent variables or the latent factors. In fact, the negative skewness computed for the three scales attested to the clustering of cases to the right of the distributions. The opposite was true for the distributions of the independent variables. Similarly, an investigation of kurtosis reveals that the distributions of the three scales were strongly peaked, with short, thick tails. Kurtosis values below zero indicated that the data associated with this index had a flat distribution, with many of the cases in the tails.

Table 41: Descriptive and distribution statistics of the three scales and continuous independent variables (n=229)

Measurement scale	Mean score	Standard deviation	Skewness		Kurtosis	
			Sk	Std. error	Ku	Std. error
Organisational professionalism	115.86	18.661	-1.241	0.161	2.316	0.320
Intrinsic Motivation	56.71	5.649	-1.392	0.161	1.506	0.320
Individual Control of Training Outcomes	22.97	3.621	-1.328	0.161	2.613	0.320
Age	41.28	11.36	0.375	0.161	-0.948	0.320
Experience (years)	20.75	11.66	0.500	0.161	-0.877	0.320
Flying time (hours)	9753.29	6116.72	0.731	0.161	-0.215	0.320
Digital time (hours)	4176.23	3216.05	0.875	0.161	0.307	0.320

Basic sample descriptive statistics, such as age, experience and flying time, are conceptually related to one another and exhibit similar distributive characteristics. Non-normality in the independent variables was expected, because, given the nature of the industry, the sample was extremely homogeneous, and all the participants had high levels of flight experience. Because the skewness and kurtosis of data do not indicate how close to normality a distribution actually is, and these indicators tend to deal with only one aspect of non-normality each (Field, 2005, 2009), to be truly useful, demographics relating to the size of the organisation had to be examined using the Kolmogorov-Smirnov and Shapiro-Wilks goodness-of-fit tests, bearing in mind the guideline that the “[s]creening of continuous variables for normality is an important early step in almost every multivariate analysis” (Tabachnick & Fidell, 2007:79).

The results of the objective statistical tests are set out in Table 42.

Table 42: Statistical tests for normality

Variables (omnibus)	Kolmogorov-Smirnov			Shapiro-Wilks			
	Statistic	df	Sig.	Statistic	df		Sig.
Age	0.098	229	0.000	0.953	229		0.000
Experience	0.117	229	0.000	0.938	229		0.000
Flying time	0.093	229	0.000	0.934	229		0.000
Digital time	0.141	229	0.000	0.918	229		0.000
Organisational Professionalism	0.096	229	0.000	0.950	229		0.000
Intrinsic Motivation	0.087	229	0.000	0.955	229		0.000
Individual Control of Training Outcomes	0.063	229	0.030	0.976	229		0.001

Variables (categorised)	Size of carrier	Mean	Kolmogorov-Smirnov			Shapiro-Wilks		
			Statistic	df	Sig.	Statistic	df	Sig.
Age (years)	Large	44	0.092	135	0.008	0.958	134	0.000
	Medium	33	0.164	48	0.002	0.858	49	0.000
	Small	41	0.112	46	0.189*	0.932	46	0.010
Experience (years)	Large	24	0.094	135	0.005	0.957	134	0.000
	Medium	13	0.176	48	0.001	0.836	49	0.000
	Small	20	0.193	46	0.000	0.870	46	0.000
Flying time (hours)	Large	11993	0.076	135	0.057	0.969	134	0.004
	Medium	5708	0.176	48	0.001	0.851	49	0.000
	Small	7537	0.188	46	0.000	0.801	46	0.000
Digital time (hours)	Large	5378	0.071	135	0.095*	0.972	134	0.008
	Medium	2663	0.225	48	0.000	0.829	49	0.000
	Small	2287	0.259	46	0.000	0.690	46	0.000
Organisational Professionalism	Large	5.86	0.096	135	0.004	0.897	134	0.000
	Medium	5.67	0.080	48	0.200*	0.972	49	0.284*
	Small	5.47	0.090	46	0.200*	0.923	46	0.005
Intrinsic Motivation	Large	5.35	0.158	135	0.000	0.886	134	0.000
	Medium	5.39	0.187	48	0.000	0.863	49	0.000
	Small	5.19	0.204	46	0.000	0.873	46	0.000
Individual Control of Training Outcomes	Large	4.33	0.118	135	0.000	0.914	134	0.000
	Medium	4.32	0.193	48	0.000	0.846	49	0.000
	Small	4.38	0.167	46	0.003	0.916	46	0.003

Significant when P < 0.05

* Not significant

The two-tailed test for significance in Table 42 shows that for the vast majority of the continuous independent variables and latent dependent factors, the distribution of the scores was statistically non-normal ($p < 0.01$).

The Shapiro-Wilks test provides a far superior computation of the test statistic (Field, 2009). Therefore, an appropriate family of non-parametric methods was used to explore phenomena in the data further, and determine the statistical significance of the relationship between the various characteristics of South African airline pilots and their perceptions of the advanced automated aircraft training climate.

5.7 RESULTS OF THE NON-PARAMETRIC COMPARATIVE STATISTICS USED TO EXPLORE PHENOMENA

The previous section has dealt with an in-depth exploratory analysis of the underlying characteristic phenomena based on the observed evidence, which manifested from within the data. However, to gain a deeper understanding of what other phenomena might be present, it was decided to conduct a series of non-parametric comparative tests. This helped the researcher, as a first step, to enhance the subsequent more complex statistical analysis in the current study. The nature of the study is primarily exploratory. In this section, non-parametric comparisons were conducted on broad categories of the independent variable. Any significance was then followed up with appropriate *post hoc* univariate examinations. Thereafter, a detailed exploration of the phenomena was conducted using a non-parametric MANOVA technique, as reported in Section 5.9 and its sub-sections.

The Kruskal-Wallis test was conducted to compare multivariate data. When data are deemed ordinal and the assumptions of the equality of group variances are violated, appropriate non-parametric statistics such as these are recommended (Field, 2005, 2009). The mean rank scores of various independent categorical groups were compared and examined to guide the more complex analyses in the study.

Where significant differences were found, a *post hoc* non-parametric test, the Mann-Whitney test, was used to determine the statistical significance of the actual difference between the highest and lowest ranking categories. When the means of the various demographic categories are tested, the most appropriate statistic used in parametric statistics is Student's t-test. However, in a case such as this, either the Mann-Whitney U or the Wilcoxon W would more be useful or appropriate as a non-parametric alternative (Green & Salkind, 2008). The Mann-Whitney test was therefore applied to the data as the equivalent to the independent sample's t-test to evaluate the differences between the medians. This test is useful in determining the significance of differences between the mean ranks for dyadic categorical independent variables and continuous dependent variables (Field, 2005; Leech *et al.*, 2005). In addition, sample sizes in excess of 40 are deemed more appropriate and large enough for this type of non-parametric comparison to yield accurate p-values (Green & Salkind, 2008). Reporting the effect of loss in statistical power may then not be necessary for Mann-Whitney comparative tests when sample sizes are sufficiently large (Field, 2009; Green & Salkind, 2008).

Although Field (2005:550) recommends applying the "Bonferroni correction" as a preventative measure against inflated Type 1 errors when conducting numerous *post hoc* tests, it was not a requirement for the present data set being examined. Therefore, combining independent variable classes to obtain sample categories that were sufficiently large and relevant for comparison mitigated the possibility of inflating Type 1 errors when conducting *post hoc* tests. Alternatively, the formula ($r = z/\sqrt{N}$) provided by Field (2005) was used as an indication of the effect size associated with any significant differences between the mean ranks of samples. It was assumed that $r \leq 0.10$ suggests no practical significance, $r > 0.10$ suggests a small effect, $r \geq 0.30$ suggests a medium effect, and $r \geq 0.50$ suggests a large effect.

The effect sizes in terms of small, medium or large effect sizes, as a method of power and practical significance analysis should, however, be read in context, because one should beware of an overly broad categorisation, based on what Ellis (2010:230) calls "T-shirt sizes". The computation of effect sizes and practical significance has been criticised by some authors as becoming increasingly generalised, which may then mask important alternative explanations or diminish useful considerations (Ellis, 2010). Thus, even where the effect size in the data seemed to diminish the impact of tests yielding significant p-values, it was considered in context for the purpose of the discussion of the nature of the phenomena observed. The concept of effect size, as originally proposed by Cohen (1988), has been adopted in the interpretation of the current research results in an effort to illuminate and complement the statistically significant findings.

The information provided in Tables 43 to 52 describes the strength in the differences or similarities between the various groups with regard to the latent behavioural factors observed in the construct. First, broad assessments on the behavioural scales in respect of the demographic groupings were made using a multivariate non-parametric comparative procedure (the Kruskal-Wallis or K-W procedure). Thereafter, a drill-down was done, using a two-sample non-parametric statistical analysis to examine differences in the mean ranks (M-W).

To examine the presence of any potentially ordered pattern in the medians of each group, the Jonckheere-Terpstra (J-T) test was used to assess the possibility of such data trends (Field, 2005, 2009). This test is useful when a researcher suspects that the order of the independent groups may be meaningful. Significant Jonckheere-Terpstra tests provide a valuable and useful basis for a more detailed understanding of the phenomena present within data. These analyses provided initial observable contact with the data set and also provided an indication of the correct path to follow for conducting further exploration. It was also then necessary to repeat the *post hoc* Mann-Whitney tests for some aspects of the data after discovering potentially important phenomena, such as Jonckheere-Terpstra trends.

Tables 43 to 52 set out the values for the chi-square and z-scores for the relevant significance tests, together with the two-tailed asymptotic levels. Some authors have argued that the one-tailed significance test cannot play an effective role in an exploratory examination of phenomena in nature (Clark & Watson, 1995). When sample sizes are relatively large, statistical significance is more likely, therefore good practice warrants an assessment of the effect associated with such significant findings (Cohen & Lea, 2004). The effect sizes were therefore generated for all *post hoc* analyses to determine the level of practicality whenever differences were detected in the data.

Table 43: Kruskal-Wallis test for the grouping variables flight deck position and size of carrier

Behavioural scale	Flight deck position	N	Mean rank	Chi-square	df	Asym. Sig. (2-tailed)
Organisational Professionalism	Co-Pilot Long-Range	39	121.46			
	Co-Pilot Short-Range	74	102.73			
	Captain Long-Range	31	121.65			
	Captain Short-Range	80	113.63			
	Total (5 missing)	224				
				3.071	3	0.381
Intrinsic Motivation	Co-Pilot Long-Range	39	104.49			
	Co-Pilot Short-Range	74	106.68			
	Captain Long-Range	31	119.32			
	Captain Short-Range	80	119.14			
	Total (5 missing)	224				
				2.393	3	0.495
Individual Control of Training Outcomes	Co-Pilot Long-Range	39	107.62			
	Co-Pilot Short-Range	74	113.97			
	Captain Long-Range	31	78.39			
	Captain Short-Range	80	126.74			
	Total (5 missing)	224				
				12.847	3	0.005*
Organisational Professionalism	Size of carrier	N	Mean rank			
	Large	135	126.37			
	Medium	48	104.33			
	Small	46	92.76			
	Total	229				
				10.416	2	0.005*
Intrinsic Motivation	Large	135	117.72			
	Medium	48	119.17			
	Small	46	102.67			
	Total	229				
				2.010	2	0.366
Individual Control of Training Outcomes	Large	135	113.48			
	Medium	48	114.44			
	Small	46	120.04			
	Total	229				
				0.341	2	0.843

*P < 0.05

Table 44: Mann-Whitney *post hoc* significance tests for the grouping variables flight deck position and size of carrier

Behavioural scale	Flight deck position	N	Mean Rank	M-W U	Z	2-tail Asymp. Sig. (effect size)
Individual Control of Training Outcomes	Captain Long-Range	31	39.98			
	Captain Short-Range	80	62.21			
	Total	111				
				743.5	-3.281	0.001* (0.31)
	All Short-Range Pilots	154	120.60			
	All Long-Range Pilots	70	94.67			
	Total (5 missing)	224				
				4 142.0	-2.791	0.005* (0.19)
	Co-Pilot Long-Range	39	40.06			
	Captain Long-Range	31	29.76			
	Total	70				
				426.5	-2.118	0.034* (0.26)
	All Co-Pilots	109	115.32			
	All Captains	120	114.71			
	Total	229				
				6 505.0	-0.070	0.944
Behavioural Scale	Size of carrier	N	Mean rank	M-W U	Z	2-tail Asymp. Sig. (effect size)
Organisational Professionalism	Larger operators	135	138.60			
	Smaller operators	94	81.71			
	Total	229				
				3202.0	-6.405	0.000* (0.42)

*P < 0.05

A non-parametric comparison of the subgroups for pilots' flight deck positions showed that the behavioural scale *Individual Control of Training Outcomes* was statistically affected by one of four flight deck positions that a pilot may occupy [$H(3) = 12.847$, $p < 0.05$]. The Mann-Whitney *post hoc* test was used to follow up on specific differences within the subgroups. The results show that short-range captains' scores were statistically higher than those of their long-range counterparts ($U = 743.50$, $p = 0.001$). It appears that the captains on short-range aircraft in the present sample felt that they had more control than the group of long-range captains over their levels of comfort in training and the related learning outcomes for the advanced aircraft they operate. The effect of this significance is regarded as medium. One possible reason for this difference may be the fact that short-range pilots have opportunities to fly more sectors than long-range pilots (therefore, they fly with a higher frequency, resulting in more exposure to their aircraft). Familiarity with an aircraft is directly related to the number of take-offs and landings performed (experience), and can therefore improve pilots' levels of confidence and comfort in working with the technology.

Similarly, there was a statistically significant difference between the mean ranked scores of *all* the short-range pilots and *all* the long-range pilots, but the effect is regarded as small ($U = 4142$, $p < 0.05$). Further examination of the long-range pilot group reveals that the co-pilots' scores were significantly higher than those of the captains ($U = 465.50$, $p < 0.05$, small effect). In general terms, captains have more overall flight experience. However, it appears that because a captain is normally a senior employee at an airline, much of a captain's flight experience is accounted for on analogue type aircraft (older generation aircraft) as opposed to digital or glass flight deck aircraft (which are more recent acquisitions at airline organisations). This difference may have affected the final behavioural scale scoring, and reduced their levels of perceived personal confidence in the advanced aircraft training they receive or have received.

Overall, pilots' perceptions on only the organisational professionalism scale were significantly affected by the size of the carrier where they were employed ($H[2] = 10.416$, $p < 0.01$). An analysis of a further *post hoc* two grouping of the large, medium and small categories (where South African Airways combined with BA Comair were considered larger operators, whilst the others were considered smaller operators, as defined in section 4.11), showed that perceptions of organisational professionalism were significantly affected (U = 3202.0, $p < 0.001$, medium effect). In addition, the order of the groupings (large, medium, small) was statistically significant (see Table 43). The Jonckheere-Terpstra test showed that this pattern was meaningful in that, as the size of the carrier increased, scores on the latent behavioural scale focused on perceptions of *Organisational Professionalism* increased statistically (J-T[3] = 4403, Std. J-T[3] = -6.036, $p < 0.001$). In other words, the larger the company in which the participant was employed, the more they would perceive their training as professional.

Table 45: Kruskal-Wallis test for the grouping variable interaction effect between experience in advanced aircraft and computer literacy

Behavioural scale	Interaction effect	N	Mean rank	Chi-square	df	2-tail Asymp.Sig.
Organisational Professionalism	Low Experience*Low Computer Literacy	19	76.00			
	Low Experience*High Computer Literacy	66	109.71			
	High Experience*Low Computer Literacy	73	130.42			
	High Experience*High Computer Literacy	71	114.49			
	Total	229				
				10.972	3	0.012*
Intrinsic Motivation	Low Experience*Low Computer Literacy	19	82.26			
	Low Experience*High Computer Literacy	66	125.74			
	High Experience*Low Computer Literacy	73	105.15			
	High Experience*High Computer Literacy	71	123.90			
	Total	229				
				2.393	3	0.025*
Individual Control of Training Outcomes	Low Experience*Low Computer Literacy	19	79.61			
	Low Experience*High Computer Literacy	66	138.09			
	High Experience*Low Computer Literacy	73	98.60			
	High Experience*High Computer Literacy	71	119.87			
	Total	229				
				18.485	3	0.000*

*P < 0.05

Table 46: Mann-Whitney *post hoc* significance tests for the grouping variable interaction effect between experience in advanced aircraft and computer literacy

Behavioural scale	Interaction effect	N	Mean rank	M-W U	Z	2-tail Asymp. Sig. (effect size)
Organisational Professionalism	Low Experience*High Computer Literacy	66	63.39			
	High Experience*Low Computer Literacy	73	75.97			
	Total	139				
				1973.0	-1.840	0.066
Intrinsic Motivation	Low Experience*High Computer Literacy	66	76.86			
	High Experience*Low Computer Literacy	73	63.80			
	Total	139				
				1956.5	-1.916	0.055
Individual Control of Training Outcomes	Low Experience*High Computer Literacy	66	82.48			
	High Experience*Low Computer Literacy	73	58.72			
	Total	139				
				1585.5	-3.492	0.000* (0.30)

*P < 0.01

Table 46 shows that overall, the effect of the interaction between a pilot's level of experience in advanced aircraft and her or his perceived computer literacy, significantly affects the pilot's perceptions of the advanced aircraft training climate ($p < 0.05$). Additionally, the mean rank scores on the Individual Control of Training Outcomes (Factor 3 score) behavioural scale is significantly different between pilots who reported low experience in advanced aircraft, combined with a high level of computer literacy (Mean Rank = 82.48) and the scores of pilots who reported high experience in advanced aircraft, combined with a low level of computer literacy (Mean Rank = 58.72).

The effect of the difference between pilots who reported low experience in advanced aircraft, combined with a high level of computer literacy, was regarded as medium ($U=1585.5$, $p < 0.001$; $r = 0.30$), and may have important consequences for airline training organisations. It appears that many hours of flight experience alone is not necessarily an important attribute in learning to operate the most advanced commercial aircraft. Organisations should consider that technologically averse individuals might have difficulty training for advanced automated aircraft.

It appears that the trainees' perceptions of their own computer literacy have a much larger influence in terms of interacting with pilots' flight experience levels at whether they feel that they may have more control of their training outcomes. More in-depth *post hoc* analysis on this interaction effect is reported in an analysis of the general linear model in Section 5.9.

Further examination of the rating on the computer literacy that the participants allocated to themselves was necessary. The results set out in Table 47 indicate that pilots' levels of computer literacy (poor, average, above average, excellent) significantly affects their *Intrinsic Motivation* and *Individual Control of Training Outcomes* regarding training for flying advanced technology aircraft (H[3] = 13.291, 19.450, $p < 0.05$). A significant difference was found between the mean ranked scores of pilots who perceived their computer literacy as low and the scores of those who perceived their computer literacy as higher (U=4961.0, 4432.5, $p < 0.01$, small to medium effect). Once again, these results show the importance of basic computer skills and competence in technology for trainees' perceptions of their training for flying advanced aircraft (see Table 47).

It was also necessary to explore whether a pattern existed in participants' perceived levels of computer literacy in terms of their learning for new technology aircraft. The Jonckheere-Terpstra test suggests that the order of computer literacy ratings (poor, average, above average, excellent) is statistically significant on the two behavioural scales at the individual level of analysis labelled *Intrinsic Motivation* and *Individual Control of Training Outcomes* (J-T[4] = 10178.50, 10888.0; Std. J-T[4] = 2.948, 4.272; $p < 0.01$).

These results attest that as pilots' perceived levels of computer literacy increase or improve, so too will their motivation to learn about new technology aircraft, and this affects their personal feelings about their ability to control the outcomes related to their training for flying advanced aircraft.

Table 47: Kruskal-Wallis test for the grouping variable computer literacy

Behavioural scale	Computer literacy	N	Mean rank	Chi-square	df	2-tail Asymp. Sig.
Organisational Professionalism	Poor	5	99.10			
	Average	87	120.34			
	Above average	92	110.92			
	Excellent	45	114.78			
	Total	229				
				1.202	3	0.752
Intrinsic Motivation	Poor	5	32.90			
	Average	87	104.30			
	Above average	92	122.76			
	Excellent	45	128.94			
	Total	229				
				13.291	3	0.004*
Individual Control of Training Outcomes	Poor	5	45.40			
	Average	87	97.51			
	Above average	92	123.25			
	Excellent	45	139.68			
	Total	229				
				19.450	3	0.000*

*P < 0.01

Table 48: Mann-Whitney *post hoc* significance tests for the grouping variable computer literacy

Behavioural scale	Computer literacy	N	Mean rank	M-W U	Z	2-tail Asymp. Sig. (effect size)
Intrinsic Motivation	Low Competence	92	100.42			
	High Competence	137	124.79			
	Total	229				
				4961.0	-2.738	0.006* (0.18)
Individual Control of Training Outcomes	Low Competence	92	94.68			
	High Competence	137	128.65			
	Total	229				
				4432.5	-3.823	0.000* (0.25)

*P < 0.01

Table 49 presents the aircraft that pilots operated according to manufacturer name. It shows that, overall, the type of aircraft operated by the pilot statistically affected their behaviour on the *Organisational Professionalism* scale and the *Individual Control of Training Outcomes* scale ($p < 0.05$). Furthermore, Table 49 clearly shows that when one compares the two largest subgroups that operate one of the main manufacturers' advanced aircraft (Boeing or Airbus), there is a significant difference in the pilots' scores on their *Individual Control of Training Outcomes* with regard to the climate for training on these aircraft ($U=2181.5$, $p < 0.05$, small effect). It appears that the Boeing pilots in the sample felt that they were more in control of their training outcomes than the Airbus pilots for this sample did. However, the effect of this difference should be regarded as small and needs to be interpreted with caution. It is recommended that this phenomenon be explored using a larger sample in future in order to obtain a more accurate effect size level.

Table 49: Kruskal-Wallis test for the grouping variable manufacturer

Behavioural scale	Manufacturer	N	Mean rank	Chi-square	df	2-tail Asymp. Sig.
Organisational Professionalism	Boeing	57	131.85			
	Airbus	95	129.53			
	Embraer	11	33.23			
	Canadair	9	87.44			
	De Havilland	7	103.93			
	Other	50	92.69			
	Total	229				
				32.454	5	0.000*
Intrinsic Motivation	Boeing	57	108.18			
	Airbus	95	116.41			
	Embraer	11	84.09			
	Canadair	9	103.50			
	De Havilland	7	119.86			
	Other	50	128.30			
	Total	229				
				5.403	5	0.369
Individual Control of Training Outcomes	Boeing	57	125.21			
	Airbus	95	104.15			
	Embraer	11	86.00			
	Canadair	9	90.39			
	De Havilland	7	139.00			
	Other	50	131.42			
	Total	229				
				11.357	5	0.045*

*$P < 0.05$

Table 50: Mann-Whitney *post hoc* significance test for the grouping variable manufacturer

Behavioural scale	Manufacturer	N	Mean rank	M-W U	Z	2-tail Asymp. Sig. (effect size)
Organisational Professionalism	Boeing	57	78.68			
	Airbus	95	75.19			
	Total	152				
				2583.5	-0.472	0.637
Individual Control of Training Outcomes	Boeing	57	85.73			
	Airbus	95	70.96			
	Total	152				
				2181.5	-2.014	0.044* (0.16)

*P < 0.05

Table 51: Kruskall-Wallis test for the grouping variables of initial (*ab initio*) training

Behavioural scale	Initial training	N	Mean rank	Chi-square	df	2-tail Asymp. Sig.
Organisational Professionalism	Military	81	127.39			
	Cadet	18	101.92			
	Self-sponsored (part-time)	64	118.45			
	Self-sponsored (full-time)	64	96.00			
	Total (2 missing)	227				
				9.083	3	0.028*
Intrinsic Motivation	Military	81	117.94			
	Cadet	18	96.86			
	Self-sponsored (part-time)	64	115.93			
	Self-sponsored (full-time)	64	111.90			
	Total (2 missing)	227				
				1.650	3	0.648
Individual Control of Training Outcomes	Military	81	109.93			
	Cadet	18	100.53			
	Self-sponsored (part time)	64	116.55			
	Self-sponsored (full-time)	64	120.39			
	Total (2 missing)	227				
				1.789	3	0.617

*P < 0.05

Table 51 and Table 52 suggest that a pilot's initial or *ab initio* flying training can significantly affect his or her perceptions of the advanced aircraft training climate at his or her airline with regard to the behavioural scale *Organisational Professionalism* [H(3) = 9.083, $p < 0.05$]. *Ab initio* training is defined as primary training, which a potential pilot undergoes when first entering the aviation industry (Moore *et al*., 2001).

The results indicate that there is a significant difference between the scores of those pilots who had military training (Mean Rank = 127.39) and the scores of pilots who had no military background (Mean Rank = 106.57) with regard to their views of *Organisational Professionalism* (U = 4828.5, $p < 0.05$).

Table 52 shows that, in terms of whether the candidate underwent a structured early training experience (that is an airline cadetship or military background) or an unstructured one (that is, participants who indicated that their primary training was concluded after self-sponsored or part-time methods), there was no impact on their perceptions of the advanced aircraft training climate with regard to the *Organisational Professionalism* scale.

It may be deduced from the results in Table 52, that having only a regimented, structured training background, as in the case of military trained pilots, may influence the candidate's perception of the professionalism associated with training for advanced aircraft. Although this information is useful, the practical significance of the differences in these categories should nevertheless be regarded as small, in terms of Cohen's (1988) criteria.

The discovery of the aforementioned effects between selected independent demographic variables and the latent factors of the main measurement construct was used as a basis for a further exploration of the possible phenomena that may exist in the dataset.

Table 52: Mann-Whitney *post hoc* significance test for the grouping variable nature of initial training

Behavioural scale	Nature of Initial training	N	Mean rank	M-W U	Z	2-tail Asymp. Sig. (effect size)
Organisational Professionalism	Structured training	99	122.76			
	Unstructured training	128	107.23			
	Total (2 missing)	227				
				5 469.0	-1.767	0.077
	Military-trained	81	127.39			
	Not military-trained	146	106.57			
	Total (2 missing)	227				
				4 828.5	-2.289	0.022* (0.15)

*P < 0.05

5.8 ASSOCIATIONAL STATISTICS: NON-PARAMETRIC MEASURES OF BIVARIATE RELATIONSHIPS

The general term explaining the concept of variable relationship is that, information about one variable is usually carried by the other variable in many instances (Cohen, Cohen, West & Alken, 2003). Exploiting the bivariate associational relationship between paired variables enhanced the exploration of the data set.

To assess the relationship or extent to which variables may be related and to determine the magnitude and subsequent direction of the possible relationship, a correlational analysis was conducted. Because the items were designed to comply with Likert's (1932) method, data were considered at least ordinal in nature. It was found that Pearson's coefficient would be unsuitable for the current study's data, as Pearson's formula is problematic where there are violations of normality or unequal variances. The difference in the non-parametric equivalent of the Pearson coefficient lies in the type of data that are used (Cohen & Lea, 2004; Comrey & Lee, 1992; Field, 2009). The non-parametric equivalent to Pearson's correlation coefficients is calculated by applying the formula to the ranks of the data, as opposed to applying it to the raw data values.

Kendall's tau-b was used as a measure of association between the variables chosen (Kendall & Stuart, 1963). Kendall's method was selected for its robustness in measuring the strength of the association between two ordinal or binary variables (Morgan *et al.*, 2007). The Kendall tau-b (τ) statistic was also used because the computation allows for adjustments for ties (here, the geometric mean is used as an estimate of the relevant tied pairs). Kendall's tau-b is equivalent to Spearman's rho in terms of the underlying assumptions, and tau "has been emphasized recently as a substitute for *r* in various research contexts" (Walker, 2003:525). However, Spearman's rho and Kendall's tau are not identical in magnitude, since their underlying logic and computational formulae are relatively different (Kendall & Stuart, 1963; Gravetter & Wallnau, 2008). Relational strengths were thus found to be more conservative under the tau statistic.

When the associational process was replicated using Spearman's formula, statistically relevant strengths of association between variables were stronger to some extent (not reported here). Nonetheless, Kendall's formula was maintained for final computations and drawing conclusions in this thesis. The relationship between two measures using Kendall's tau formula is based on the goodness-of-fit of the least squares straight line; however, this correlation simply provides some evidence of a relationship and will therefore never prove causality *per se* (Tabachnick & Fidell, 2007). Cohen and Lea (2004) suggest a cut-off point of 0.30 (a medium effect size) for the practical significance of association.

The three main components of the construct under study were correlated with some core demographic data (see Table 53).

Table 53: Main demographic and factor correlations

Kendall's tau-b	Flight deck position	Interact_ Group	Adv. Aircraft Exp.	Age	Gender	Level of education	Level of computer literacy	Pilot unionisa-tion	Size of carrier	Instruc-tor rated	F1	F2	F3
Flt deck position	1.000	0.060	0.066	0.328**	0.128*	0.145*	0.000	0.176**	0.216**	0.057	0.003	0.075	0.068
Sig (2-tailed)	.	0.291	0.283	0.000	0.038	0.019	0.995	0.004	0.000	0.353	0.949	0.151	0.199
Interact_Group	0.060	1.000	0.810**	0.293**	0.060	-0.006	0.193**	-0.348**	-0.364**	0.069	0.082	0.053	-0.004
Sig (2-tailed)	0.291	.	0.000	0.000	0.325	0.918	0.002	0.000	0.000	0.262	0.104	0.307	0.941
Adv. Aircraft Exp.	0.066	0.810**	1.000	0.413**	0.058	-0.011	-0.279**	-0.442**	-0.452**	0.039	0.123*	-0.010	-0.100
Sig (2-tailed)	0.283	0.000	.	0.000	0.379	0.863	0.000	0.000	0.000	0.557	0.024	0.857	0.077
Age	0.328**	0.293**	0.413**	1.000	0.206**	0.122	-0.211**	-0.210**	-0.192**	0.035	0.143**	0.016	-0.050
Sig (2-tailed)	0.000	0.000	0.000	.	0.002	0.065	0.001	0.002	0.002	0.596	0.009	0.781	0.372
Gender	0.128*	0.060	0.058	0.206**	1.000	0.009	0.040	-0.002	0.045	0.048	0.072	0.097	0.038
Sig (2-tailed)	0.038	0.325	0.379	0.002	.	0.889	0.548	0.974	0.472	0.470	0.185	0.083	0.507
Level of education	0.145*	-0.006	-0.011	0.122	0.009	1.000	-0.011	0.033	-0.013	0.154*	0.031	0.046	0.037
Sig (2-tailed)	0.019	0.918	0.863	0.065	0.889	.	0.864	0.622	0.839	0.020	0.576	0.409	0.513
Level of com, literacy	0.000	0.193**	-0.279**	-0.211**	0.040	-0.011	1.000	0.133*	0.126*	0.054	-0.043	0.153**	0.216* *
Sig (2-tailed)	0.995	0.002	0.000	0.001	0.548	0.864	.	0.044	0.047	0.414	0.433	0.006	0.000
Pilot unionisation	0.176**	-0.348**	-0.442**	-0.210**	-0.002	0.033	0.133*	1.000	0.854**	-0.025	-0.319**	-0.068	-0.010
Sig (2-tailed)	0.004	0.000	0.000	0.002	0.974	0.622	0.044	.	0.000	0.704	0.000	0.222	0.862
Size of carrier	0.216**	-0.364**	0.452**	-0.192**	0.045	-0.013	0.126*	0.854**	1.000	-0.012	-0.315**	-0.091	-0.027
Sig (2-tailed)	0.000	0.000	0.000	0.002	0.472	0.839	0.047	0.000	.	0.843	0.000	0.089	0.614

Kendall's tau-b	Flight deck position	Interact_ Group	Adv. Aircraft Exp.	Age	Gender	Level of education	Level of computer literacy	Pilot unionisa-tion	Size of carrier	Instruc-tor rated	F1	F2	F3
Instructor rated	0.057	0.069	0.039	0.035	0.048	0.154*	0.054	-0.025	-0.012	1.000	-0.023	0.050	0.111
Sig (2-tailed)	0.353	0.262	0.557	0.596	0.470	0.020	0.414	0.704	0.843	.	0.677	0.370	0.050
Organisational Professionalism (F1)	0.003	0.082	0.123*	0.143**	0.072	0.031	-0.043	-0.319**	-0.315**	-0.023	1.000	0.351**	0.448**
Sig (2-tailed)	0.949	0.104	0.024	0.009	0.185	0.576	0.433	0.000	0.000	0.677	.	0.000	0.000
Intrinsic Motivation (F2)	0.075	0.053	-0.010	0.016	0.097	0.046	0.153**	-0.068	-0.091	0.050	0.351**	1.000	0.396**
Sig (2-tailed)	0.151	0.307	0.857	0.781	0.083	0.409	0.006	0.222	0.089	0.370	0.000	.	0.000
Individual Control of Training Outcomes (F3)	0.068	-0.004	-0.100	-0.050	0.038	0.037	0.216**	-0.010	-0.027	0.111	0.448**	0.396**	1.000
Sig (2-tailed)	0.199	0.941	0.077	0.372	0.507	0.513	0.000	0.862	0.614	0.050	0.000	0.000	.

** $p < 0.001$;
* $p < 0.05$;
$\tau_s < 0.10$ suggests no effect;
$\tau_s \geq 0.10$ suggests a small effect;
$\tau_s \geq 0.30$ suggests a medium effect; and
$\tau_s \geq 0.50$ suggests a large effect

The correlation results depicted in Table 53 indicate that, for the present sample, the phenomena discussed below were noteworthy.

The position of the crewmember on the flight deck (either captain or co-pilot) is significantly correlated with age ($\tau_s = 0.328$, $p < 0.001$, medium effect). Due to the very rigid seniority systems entrenched at the larger and unionised carriers in South Africa, a pilot is only eligible for command (a captain upgrade) after serving a prerequisite number of years in the organisation. The data therefore reflects this correlation. The results also show a small effect size in the relationship between a respondent's flight deck position and his or her level of education ($\tau_s = 0.328$, $p < 0.001$). The data indicated that it is to be expected that, as a pilot gains experience and becomes more senior at an airline, the pilot's educational qualifications will improve.

The interaction effects between pilots' overall experience level on advanced aircraft and their perceived levels of computer literacy are significantly related to the ages of the pilots ($\tau_s = 0.293$, $p < 0.001$, medium effect). This correlation suggests that levels of the experience-computer literacy interaction for this sample improved with the age of the pilots. It may therefore be relevant that older pilots who undergo advanced aircraft transition training have a vast amount of previous experience in advanced aircraft, together with better perceived computer literacy, and that this tends to improve their training experience. Similarly, it is intuitively logical to conclude that as people spend more years in an airline, they both age, and are likely to gain experience. However, levels of computer literacy do not display the same linear relationship. Therefore, older pilots may require higher levels of computer literacy or ability in order to report an improved perception of the advanced aircraft training they undergo.

The level of advanced aircraft flight experience of pilots in the sample was inversely related to their level of computer literacy ($\tau_s = -0.279$, $p < 0.001$, medium effect), the degree to which the pilot group was unionised ($\tau_s = -0.442$, $p < 0.001$, medium effect) and the size of the organisation ($\tau_s = -0.452$, $p < 0.001$, medium effect). A surprising correlation in the data suggests that the more experience a pilot had in advanced aircraft, the less likely the pilot was to believe that he or she had good levels of computer literacy. This result may also be related to the effect of age on perceived levels of computer literacy, in that younger pilots tended to have more positive perceptions of their computer literacy ($\tau_s = -0.211$, $p < 0.001$, small to medium effect). It appears that the older pilots (over the age of 40 years) have had more experience on advanced aircraft ($\tau_s = 0.413$, $p < 0.001$), but less experience with commercial technology. This particular generation did not have as much exposure to computer-based technology while growing up as the younger generation has had (Moore, 2003). This effect is also related to a respondent's number of years of employment at the organisation.

Higher levels of computer literacy were significantly associated with greater pilot unionisation ($\tau_s = 0.133$, $p < 0.05$, small effect) and organisational size ($\tau_s = 0.126$, $p < 0.05$). These results suggest that the larger, unionised carriers employ pilots who perceive their levels of computer literacy to be relatively good. The results may also be related to the fact that one large unionised carrier in the sample issues personal laptop computers to its pilots. Such a pilot's familiarity with this item may then influence the pilot's perception of his or her computer literacy. Another explanation may be that, because pilots at highly unionised carriers command more earnings from complex negotiated agreements (Olney, 1996), they are more likely to be able to afford and enjoy the latest technologically advanced personal gadgets, such as laptops and tablet computers. Thus, these pilots are less likely to be averse to technology.

Greater unionisation of airline pilots is significantly related to high experience levels in advanced aircraft ($\tau_s = 0.442$, $p < 0.001$, medium effect). This was expected, as the large carriers generally hire the most experienced pilots in the industry. Unionised carriers are also more attractive to experienced pilots because of their strict seniority lists and benefits such as pensions and protection against a loss of their licences (Olney, 1996). Larger carriers in South Africa are known to operate the more advanced technology aircraft from the two main global aircraft manufacturers, Airbus and Boeing, which produce larger aircraft, flying longer distances. Pilots at the bigger, unionised carriers also fly higher frequency schedules, providing their pilots with an opportunity to gain more experience. Consequently, it was observed that in South Africa highly unionised carriers are significantly related to larger employers (τ_s

$= 0.854$, $p < 0.001$, large effect). These effects have important implications for smaller, non-unionised organisations. Such organisations are more likely to have inferior advanced aircraft training capability in terms of the perceptual behavioural scales, and therefore more effort should be made by these companies' management to ensure effective and efficient transfer of knowledge to their pilots, especially because inadequate training paradigms, structure and methodology can have an adverse impact on flight safety.

The latent factor *Organisational Professionalism* was positively related to a respondent's experience in advanced aircraft ($\tau_s = 0.123$, $p < 0.05$) and age ($\tau_s = 0.143$, $p < 0.001$). However, the effect size of this relationship was considered small. Similarly, the latent construct *Intrinsic Motivation* was positively associated with a pilot's level of computer literacy ($\tau_s = 0.153$, $p < 0.001$, small effect), which suggests that when these airline pilots perceive their computer literacy levels as improving, they may also have a greater interest in training for new technology. Furthermore, the third latent construct, *Individual Control of Training Outcomes*, was similarly correlated with pilots' perceptions of their computer literacy ($\tau_s = 0.216$, $p < 0.001$, small effect). This result indicates that as pilots begin to believe that their levels of computer literacy are relatively good, so too will their perceptions of their ability to be in control of, and to take charge of their advanced aircraft training outcomes.

The correlation matrix in Table 53 also shows that the three latent behavioural scales of the main measurement construct correlate with each other to a high degree ($p < 0.001$). This was expected because the construct was developed in terms of the systemic principle and factors derived from an oblique rotation. The tau inter-correlation coefficients for the factors ranged from 0.351 to 0.448.

Overall, the results of the non-parametric associational analysis should be interpreted within the present context and with a degree of caution. The non-parametric methods are considered more robust than parametric methods. Nonetheless, it was observed there were many low to medium effect sizes, according to Cohen's (1988) criteria, representing the overall effectiveness of significant associations. This indicates that many of the significant results can be dismissed from a practicality perspective. Again, however, effect sizes may also bear the brunt of a certain level of criticism for being subjectively broad in terms of unilaterally lessening the impact of a study, which Ellis (2010:230) dismisses as categorising significances according to “T-shirt sizes”. Nevertheless, it was necessary to understand why the data yielded such low to medium effect sizes. It was summarily hypothesised then, that one reason for this result might be from the impact of the relatively tiny and highly homogeneous sample frame and subsequent non-normality of the data set. These results pointed to areas for further exploration. Because of the lack of total clarity from this associational analysis, the aforementioned results were then used to guide further multivariate and regression statistical analyses, as discussed in the sections below.

5.9 NON-PARAMETRIC MULTIVARIATE ANALYSIS OF VARIANCE (MANOVA)

“Generalized linear models provide a unified theoretical and conceptual framework for many of the most commonly used statistical methods” (Dobson & Barnett, 2008:15). To examine the main and interactional effects of partially independent categorical variables on multiple dependent variables, a MANOVA was conducted by means of the general linear model (Anderson, 2001; Field, 2005). In this case, the dependent variables (factors) were conceptually related to a high degree. The risk of multicollinearity should be considered when correlations between the dependent variables are generally high, while conversely no correlation would imply that a multivariate analysis could be “pointless” (Leech *et al.*, 2005:177). A moderate median ranked correlation of 0.40 between the dependent variables was calculated, based on Kendall’s tau-b from an earlier associational exploration (τ, which makes adjustments for ties). This indicated that an analysis of a general linear model was indeed useful.

Based on a combination of the dependent variables, the general linear model procedure was used to compute a multivariate F. “The larger the value of F, the more likely it is that the null hypothesis (H_0) of no differences among the group means (locations) is false” (Anderson, 2001:34). The combination maximised the differentiation of the ordinal dependent variable groups. This procedure was followed to provide an analysis for *effects* on a linear combination of three dependent variables of multiple independent variables, or covariates. A MANOVA was computed to test the differences in the centroid or vector of medians of the multiple interval/ordinal dependent variables, for various categories of the independent variables. Because the Type I error rate of the standard MANOVA test statistics can be inflated, whereas their power attenuates when assumptions of normality and homogeneous covariance matrices are violated (Holmes, 2005), a non-parametric MANOVA was computed. Both Box's M-test (p = 0.001), which was used to assess the homogeneity of the variance-covariance matrices of the independent variables, and Levene's test of the equality of error variances in two of the three dependent variables were markedly violated (see Table 54).

Furthermore, assumptions on the dependent variable had to be considered as deviating from normality, because in a MANOVA, there is no single dependent variable as such, but rather a column matrix or vector of scores on each dependent variable.

Table 54: Tests for assumptions of normality and homogeneity

Levene's test of equality of error variances				
Ranked dependent factor or scale	F	df1	df2	Sig.
Organisational Professionalism (Factor 1)	1.462	51	177	0.037
Intrinsic Motivation (Factor 2)	1.938	51	177	0.001
Individual Control of Training Outcomes (Factor 3)	1.281	51	177	0.122
Box's M-test				
273.269	1.426	138	3699.387	0.001

In this situation (as portrayed in Table 54), a non-parametric or rank-order variant of the MANOVA was an appropriate option, as proposed by Zwick (1985). Original or raw observations were transformed by ranking the subjects on each of the dependent variables. The ranks were then subjected to a conventional MANOVA. However, the required test statistic in this case was equal to (N-1)V, where V is the Pillai-Bartlett statistic computed on the transformed data (Zwick, 1985). Five independent variables based on a previous theoretical premise and earlier associational analysis, together with appropriate strengths of association (phi), were selected and tested. A breakdown of the frequencies in terms of the independent categories is provided in Table 55.

To guide an examination and interpret the differences between the vectors of the mean ranked scores between groups, the following situation was subsequently investigated:
What are the phenomena that affect airline pilots who differ in

- *age* (40 years and younger, or over 40 years);
- *level of digital flight time experience* (where high experience meant 2001 hours or more on advanced aircraft, and low experience meant 2000 hours or less);
- *company status* (captain or co-pilot);
- *size of carrier* (employed at a large, medium or small company); and
- *level of computer literacy* (poor, average, above average, excellent)

on some linear combination of the three dependent factors (*Organisational Professionalism*, *Intrinsic Motivation* and *Individual Control of Training Outcomes*). In addition, interaction between the demographic levels in a distinguishing linear combination of the dependent variables was examined. The number of respondents in each category was collapsed and is depicted in Table 55.

Table 55: Frequency of between-subjects factors

Independent demographic grouping	Sub-grouping	Valid N
Age category	40 and under	119
	41 and over	110
Level of digital flight time experience	High digital	144
	Low digital	85
Company status	Captain	120
	Co-pilot	109
Size of carrier	Large	135
	Medium	48
	Small	46
Computer literacy	Poor	5
	Average	87
	Above average	92
	Excellent	45

The results of the non-parametric MANOVA of the five demographic variables in terms of the respondents' perceptions of the training climate associated with advanced aircraft are set out in Table 56.

Table 56: Omnibus Pillai-Bartlett multivariate test of significance

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial eta squared	Observed power
Intercept	Pillai-Bartlett Trace	0.618	94.182	3	175	0.000	0.618	1.000
Age category	Pillai-Bartlett Trace	0.012	0.721	3	175	0.540	0.012	0.202
Digital flight experience	Pillai-Bartlett Trace	0.026	1.533	3	175	0.208	0.026	0.400
Company status	Pillai-Bartlett Trace	0.012	0.681	3	175	0.565	0.012	0.192
Size of carrier	Pillai-Bartlett Trace	0.194	6.311	3	352	0.000	0.097	0.999
Computer literacy	Pillai-Bartlett Trace	0.125	2.555	3	531	0.007	0.042	0.940
Digital flight experience* Computer literacy	Pillai-Bartlett Trace	0.149	3.090	3	531	0.001	0.050	0.976

The multivariate Pillai-Bartlett table tests the hypothesis that airline pilots, based on the selected demographics, do not differ significantly in terms of their overall perception of the advanced aircraft training climate. The result of the MANOVA depicted in Table 56 indicates that age of the respondents (F = 0.721; p = 0.540), digital flight experience (F = 1.533; p = 0.208) and whether the respondent was a captain or first officer (F = 0.681; p = 0.565) had no noticeable effect.

Then again, the *size of the carrier* in which the respondent is employed, the person's *level of computer literacy* and *digital flight experience***computer literacy* interaction have a substantive effect on the respondents' perceptions. For all three effects, the observed significance level for the Pillai-Bartlett test was at a 0.01 level of significance.

It appears that the size of the carrier (company) is possibly the most important independent variable in the model, combined with a high, observed power, indicating that the chance of failing to detect an effect that is present is less than a 1% (Arrindell & Van der Ender, 1985; Cohen, 1988). The Pillai-Bartlett trace was equal to 0.194, with an associated F (3, 229) = 6.311, $p < 0.001$ (Table 56). The squared eta of 0.097 indicates that the size of the carrier explained almost 10% of the variance in the specified model. The chi-square test statistic furthermore confirmed the inequality of the location of the median ranked scores of the three subgroups for the size of the company [χ^2 (3) = 1438.908, $p < 0.001$].

The effect of respondents' *digital flight experience*computer literacy* interaction (the multivariate Pillai-Bartlett's trace was 0.149; F [3, 229] = 3.090; $p < 0.001$) was significant, with a high observed power, providing only a 2% chance of failing to detect an effect which exists (Arrindell & Van der Ender, 1985; Cohen, 1988, Zwick, 1985). This indicates that the practical implications of subjects' experience in advanced aircraft with regard to their perception of the advanced aircraft training climate was highly dependent on their level of computer literacy.

The chi-square test statistic also confirmed the inequality of the localities of the median ranked scores for the following subgroups of the *digital flight experience*computer literacy* interaction [χ^2 (3) = 33.972, $p < 0.001$]:

- low advanced aircraft experience and poor, or average, or above average or excellent computer literacy (four subgroups); and
- high advanced aircraft experience and poor, or average, or above average or excellent computer literacy (four subgroups).

Overall, it appears that airline pilots' perceived levels of computer literacy had a significant effect on their experiences with the advanced aircraft training climate (multivariate Pillai-Bartlett's trace = 0.125; F [3, 229] = 2.555; $p < 0.01$). In addition, the chi-square test statistic also confirmed the inequality of the location of the median ranked scores for the four subgroups of computer literacy [χ^2 (3) = 28.5, $p < 0.001$].

5.9.1 Between-subjects effects

To determine exactly where the variations in the median ranked centroids were for each of the subgroups, across the three latent behavioural factors (*Organisational Professionalism*, *Intrinsic Motivation*, and *Individual Control of Training Outcomes*), it was necessary to proceed with an analysis of between-subjects effects. The examination was based on the results of the significance tests, and effect sizes were determined using Cohen's (1988) criterion of the partial eta squared (η^2). The recommendation for practical significance, based on this computation, is that there is a small effect size when $\eta^2 = 0.01$ (1%), a medium effect size when $\eta^2 = 0.06$ (6%) and a large effect size when $\eta^2 = 0.15$ (15%). The between-subjects effects was only examined for significant differences with regard to the size of the carrier, computer literacy and the interactive effect of digital flight experience combined with subjects' levels of computer literacy. For the size of the carrier, the results show that respondents' perceptions of the advanced aircraft training climate was significant ($p < 0.001$ with a medium effect) only for the *Organisational Professionalism* behavioural scale. The size of the carrier in which the pilot is employed appears to account for 13.9% of the variability in the pilot's perception of the professionalism in training for advanced aircraft. No statistically significant differences ($p = 0.724$) were noted between airline pilots' levels of computer literacy and their median ranked scores on the *Organisational Professionalism* scale. However, trainee's level of computer literacy had a small to medium effect ($\eta^2 = 0.047$ to 0.072) on the trainee's judgement on the *Intrinsic Motivation* and *Individual Control of Training Outcomes* scales respectively. Therefore, approximately 7.2% of the variance in the airline pilots' perceptions of their ability to control an outcome of training for advanced aircraft was related to their perceived level of computer literacy. These results show that information regarding computer skill and competence can be very useful in practical situations, such as airline pilot recruitment. To a recruiter, this may indicate a candidate's understanding, easiness and aptitude to operate advanced technology, which may in turn enhance a pilot's ability to learn successfully, and confidently operate the modern digital aircraft.

Table 57: Significance tests for between-subjects effects for Factors 1, 2 and 3

Source	Dependent variable	df	Mean square	F	Sig.	Partial eta squared	Observed power	Effect size (Cohen, 1988)
Corrected Model	*Organisational Professionalism*	51	8 661.756	2.746	0.000	0.442	1.000	Large
	Intrinsic Motivation	51	6 932.425	1.916	0.001	0.356	1.000	Large
	Individual Control of Training Outcomes	51	6 991.335	1.952	0.001	0.360	1.000	Large
Intercept	*Organisational Professionalism*	1	51 0716.861	161.885	0.000	0.478	1.000	Large
	Intrinsic Motivation	1	69 4995.446	192.101	0.000	0.520	1.000	Large
	Individual Control of Training Outcomes	1	71 9711.892	200.957	0.000	0.532	1.000	Large
Size of carrier	*Organisational Professionalism*	2	45 077.610	14.289	0.000	0.139	0.999	Medium
	Intrinsic Motivation	2	333.460	0.092	0.912	0.001	0.064	Insig.
	Individual Control of Training Outcomes	2	1 827.355	0.510	0.601	0.006	0.133	Insig.
Computer literacy	*Organisational Professionalism*	3	1 391.231	0.441	0.724	0.007	0.137	Insig.
	Intrinsic Motivation	3	10 429.173	2.883	.037	0.047	0.681	Small
	Individual Control of Training Outcomes	3	16 458.487	4.596	0.004	0.072	0.884	Medium
Level of digital flight time experience * Computer literacy	*Organisational Professionalism*	3	11 045.181	3.501	0.017	0.056	0.773	Small
	Intrinsic Motivation	3	19 322.494	5.341	0.002	0.083	0.929	Medium
	Individual Control of Training Outcomes	3	26 389.302	7.368	0.000	0.111	0.984	Medium

Finally, Table 57 shows that the level of *interaction* between Digital flight time experience and Computer literacy, significantly affects pilots' perceptions of the training climate associated with advanced aircraft training. The analysis of between-subjects effects show a significant relationship (p = 0.05 to 0.001) between Digital flight experience*Computer literacy interaction for all three latent factors in the measurement construct with a small to medium effect, according to Cohen's (1988) criterion.

These results show that the respondent's perceptions on the *Organisational Professionalism*, *Intrinsic Motivation* and *Individual Control of Training Outcomes* behavioural scales depend strongly on the joint effect of the participant's advanced aircraft flight experience, combined with his or her levels of computer competence. In other words, the two variables independently may have very little impact on perceptions of the training climate associated with the advanced aircraft; however, taken in unison, they appear to interact in such a way that this joint effect significantly separates categorical groupings on the various perception scales.

5.9.2 Non-parametric comparative *post hoc* tests for independent samples (Mann-Whitney) based on the GLM results

An omnibus F-test indicated only that the centroids of the median ranked scores were not co-located (Tabachnick & Fidell, 2007). Therefore, a series of *post hoc* non-parametric comparisons were carried out between bivariate subgroups from each category to ascertain the exact location of significant differences. Because the distributions of the scores on the dependent variables were regarded as non-normal, the Mann-Whitney test was once more used to determine any significant differences. The method is ideal for unequal and small sample sizes (Babbie, 2010), as in the case of the computer literacy subgroups. The z-value provided by the test is an indication of whether the two subgroups come from the same distribution. However, other common interpretations of the Mann-Whitney non-parametric procedure are that the test actually checks the median equality of the two samples (Field, 2005). Table 58 depicts the calculated z-values and effect size for the three categories in terms of the size of the carrier, and four categories related to the respondent's level of computer literacy. Table 59 in turn depicts the results for the calculated z-values and effect size for the six categories of Digital flight time experience* Level of computer literacy.

5.9.3 Size of the carrier

Results from the non-parametric test (see Table 58) indicate that there is a significant difference between the perceptions of pilots employed at either a small organisation or a large organisation ($p < 0.01$). When one compares the large airline pilots’ scores to those of the pilots from the group of smaller airline pilots, the scores of the pilots from the larger carriers are significantly higher only on the *Organisational Professionalism* scale (at a macro and intermediate level of the main measurement construct; $Z = -2.941$, $p < 0.01$, with a medium effect size, $r = 0.396$). The results indicate that, it is likely that pilots employed at the relatively larger airlines will have a better perception of the training and organisational structures that are in place with regard to the training they experience for advanced aircraft. According to Child (1973), “Size of organization has often been cited as the attribute having the greatest single influence on the extent to which organizations develop bureaucratic forms of organization structure”. In addition, larger enterprises have access to far more resources and greater budgets, thus creating an impression of professionalism. Another hypothesised reason for this phenomenon may be that because there are more opportunities for trainees to come into contact with flight instructors and management outside of training events at smaller airline companies, it may begin to create a casual atmosphere for learning; by virtue of over-familiarity, perceptions of professionalism can be reduced (Katz & Khan, 1966).

The medium effect size of the difference between larger and smaller carriers implies that it is of practical importance to understand the phenomenon of organisational professionalism and its subsequent impact in designing and developing training interventions and methodologies. In addition restructuring specific areas of the training may be necessary to enhance perceptions of training professionalism. For relatively small operators, it may be prudent to ensure that the training centre is well organised in terms of available learning resources, and that pertinent information about training issues (such as programmes, the names of instructors and timetables) are communicated effectively and timeously.

Table 58: Non-parametric comparison of mean rank scores by size of carrier and level of computer literacy

Latent behavioural scale	Variable	N	Mean rank score	U	Z	Asym. Sig. two-tailed	Effect size *r*	
Size of carrier								
Organisational Professionalism	Large Small Total	135 46 181	97.68 71.39	2203.00	-2.941	0.003*	0.396	Medium
Organisational Professionalism	Medium Large Total	48 135 183	78.81 96.69	2607.00	-2.009	0.440	-	-
Computer Literacy								
Intrinsic Motivation	Poor Average Total	5 87 92	16.20 48.24	66.00	-2.617	0.009*	0.273	Small
Individual Control of Training Outcomes	Poor Average Total	5 87 92	24.50 47.76	107.50	-1.905	0.057	-	-
Intrinsic Motivation	Poor Above Average Total	5 92 97	14.70 50.86	58.50	-2.810	0.005*	0.285	Small
Individual Control of Training Outcomes	Poor Above Average Total	5 92 97	18.00 50.68	75.00	-2.547	0.011*	0.259	Small
Intrinsic Motivation	Poor Excellent Total	5 45 50	8.00 27.44	25.00	-2.848	0.004*	0.403	Medium
Individual Control of Training Outcomes	Poor Excellent Total	5 45 50	8.90 27.34	29.50	-2.706	0.007*	0.383	Medium
Intrinsic Motivation	Average Above Average Total	87 92 179	82.66 96.94	3363.50	-1.849	0.064	-	-
Individual Control of Training Outcomes	Average Above Average Total	87 92 179	79.39 100.03	3079.00	-2.678	0.007*	0.200	Small
Intrinsic Motivation	Average Excellent Total	87 45 132	61.40 76.36	1514.00	-2.137	0.033*	0.186	Small
Individual Control of Training Outcomes	Average Excellent Total	87 45 132	58.36 82.24	1249.00	-3.418	0.001*	0.297	Small
Intrinsic Motivation	Above Average Excellent Total	92 45 137	67.95 71.14	1973.50	-0.445	0.657	-	-
Individual Control of Training Outcomes	Above Average Excellent Total	92 45 137	65.53 76.09	1751.00	-1.473	0.141	-	-

- P < 0.05

5.9.4 Computer literacy

Table 58 shows that there are statistically significant differences ($p \leq 0.05$ to $p < 0.01$) between the pilots who regarded their level of computer literacy as either poor, average, above average or excellent, on most of the latent scales of the research construct. The more prominent differences appeared between participants who regarded their competence in computers as poor to those who felt that their competence was excellent.

Airline pilots who regarded their level of computer literacy as poor were more negative about their training experiences for advanced aircraft, with regard to the microelements of the construct (that is, at an individual level of analysis). In other words, these participants felt less motivated to learn about advanced aircraft ($Z = -2.848$; $p < 0.01$, $r = 0.403$) and felt a sense of a loss in control over the outcomes of such training ($Z = -2.706$; $p < 0.01$, $r = 0.383$).

Furthermore, it appears from the results that when the respondents reported an improvement in their computer literacy skills, their perception of their ability to take charge of their learning for advanced technology aircraft also improved at a statistically significant level. Less significant differences appeared, however, at the upper echelon of the computer literacy band (that is, differences between above average and excellent computer skills). This may be regarded as a ceiling effect.

5.9.1 Digital flight time experience* Level of computer literacy

The aforementioned results indicated that there were significant differences between the effects of the interaction between pilots' experience in advanced technology aircraft combined with their level of computer literacy across most of the demographic categories and the latent scales. This effect warranted further investigation.

According to the results depicted in Table 57 the respondents' perceptions on the latent behavioural scales *Organisational Professionalism*, *Intrinsic Motivation* and

Individual Control of Training Outcomes were significantly ($p = 0.05$ to $p < 0.01$) affected by this interaction, with a small to medium effect size in terms of Cohen's (1988) criterion.

A closer examination of the effect of the interaction between pilots' digital flight experience combined with their level of computer literacy in Table 59 revealed that the impact of subjects' advanced aircraft experience levels on their perceptions of the research construct depended more on their perceived computer literacy levels than on any other variable (such as size of carrier or age). In general it seems that among pilots with a low level of digital flying experience, those who felt that they had a high computer literacy tended to score higher on the three behavioural scales ($p = 0.05$ to $p < 0.01$). The results furthermore suggest that airline pilots with high levels of computer literacy and high digital flight time experience were statistically ($p < 0.05$) more positive about the overall advanced aircraft training climate than the respondents in all the other categories. In other words, it appears that computer literacy is a significant intervening variable on the potential impact that actual digital flight experience has on the training attitude of pilots. It may be posited at this point that low perceived computer literacy is linked to technological averseness, and that higher perceived computer literacy levels may substitute for digital flight experience.

The overall results of the *post hoc* analyses suggest that the combined effect of a pilot's level of computer literacy and experience levels on advanced aircraft have an important role to play in creating an understanding of the phenomena associated with pilots' perceptions of the training climate. It appears that high levels of computer literacy compensate for a lack of experience on the aircraft, in respect of the pilots' confidence and the way respondents feel and experience their training for new technology aircraft. Although the effect size was small in all significant differences of the interaction effect categories, the finding may be an important contribution to the current theory (see Table 59).

Table 59: Non-parametric comparison of the mean rank scores by the level of Digital flight time experience*Computer literacy

Latent perception scale	Level of digital flight time experience* Computer literacy	N	Mean rank score	U	Z	Sig.	Effect size *r*	
Organisational Professionalism (F1)	Low Experience*Low Computer Literacy	19	32.05	419.0	-2.195	0.028	0.238	Small
	Low Experience*High Computer Literacy	66	46.15					
	Total	85						
Intrinsic Motivation (F2)	Low Experience*Low Computer Literacy	19	31.29	404.5	-2.357	0.018	0.256	Small
	Low Experience*High Computer Literacy	66	46.37					
	Total	85						
Individual Control of Training Outcomes (F3)	Low Experience*Low Computer Literacy	19	27.18	326.5	-3.194	0.001	0.346	Medium
	Low Experience*High Computer Literacy	66	47.55					
	Total	85						
Organisational Professionalism (F1)	Low Experience*Low Computer Literacy	19	31.13	401.5	-2.818	0.005	0.294	Small
	High Experience*Low Computer Literacy	73	50.50					
	Total	92						
Intrinsic Motivation (F2)	Low Experience*Low Computer Literacy	19	38.00	532.0	-1.562	0.118	-	-
	High Experience*Low Computer Literacy	73	48.71					
	Total	92						
Individual Control of Training Outcomes (F3)	Low Experience*Low Computer Literacy	19	39.32	557.0	-1.324	0.186	-	-
	High Experience*Low Computer Literacy	73	48.37					
	Total	92						
Organisational Professionalism (F1)	Low Experience*Low Computer Literacy	19	32.82	433.5	-2.384	0.017	0.251	Small
	High Experience*High Computer Literacy	71	48.89					
	Total	90						

Latent perception scale	Level of digital flight time experience* Computer literacy	N	Mean rank score	U	Z	Sig.	Effect size *r*	
Intrinsic Motivation (F2)	Low Experience*Low Computer Literacy	19	32.97	436.5	-2.365	0.018	0.249	Small
	High Experience*High Computer Literacy	71	48.85					
	Total	90						
Individual Control of Training Outcomes (F3)	Low Experience*Low Computer Literacy	19	33.11	439.0	-2.342	0.019	0.247	Small
	High Experience*High Computer Literacy	71	48.82					
	Total	90						
Organisational Professionalism (F1)	Low Experience* High Computer Literacy	66	63.39	1 973.0	-1.840	0.066	-	-
	High Experience* Low Computer Literacy	73	75.97					
	Total	139						
Intrinsic Motivation (F2)	Low Experience* High Computer Literacy	66	76.86	1 956.5	-1.916	0.055	-	-
	High Experience* Low Computer Literacy	73	63.80					
	Total	139						
Individual Control of Training Outcomes (F3)	Low Experience* High Computer Literacy	66	82.48	1 585.5	-3.492	0.000	0.296	Small
	High Experience* Low Computer Literacy	73	58.72					
	Total	139						
Organisational Professionalism (F1)	Low Experience* High Computer Literacy	66	67.17	2 222.0	-0.521	0.602	-	-
	High Experience* High Computer Literacy	71	70.70					
	Total	137						
Intrinsic Motivation (F2)	Low Experience* High Computer Literacy	66	69.52	2 309.0	-0.147	0.883	-	-
	High Experience* High Computer Literacy	71	68.52					
	Total	137						

Latent perception scale	Level of digital flight time experience* Computer literacy	N	Mean rank score	U	Z	Sig.	Effect size *r*	
Individual Control of Training Outcomes (F3)	Low Experience* High Computer Literacy	66	75.06	1 943.0	-1.736	0.083	-	
	High Experience* High Computer Literacy	71	63.37					
	Total	137						
Organisational Professionalism (F1)	High Experience* Low Computer Literacy	73	77.95	2 193.5	-1.591	0.112	-	-
	High Experience* High Computer Literacy	71	66.89					
	Total	144						
Intrinsic Motivation (F2)	High Experience* Low Computer Literacy	73	66.64	2 163.5	-1.717	0.086	-	-
	High Experience* High Computer Literacy	71	78.53					
	Total	144						
Individual Control of Training Outcomes (F3)	High Experience* Low Computer Literacy	73	65.51	2 081.5	-2.049	0040	0.171	Small
	High Experience* High Computer Literacy	71	79.68					
	Total	144						

5.10 BACKWARD STEPWISE LOGISTIC REGRESSION

Logistic regression was used as a predictive analysis to classify subjects with positive perceptions regarding their training for advanced technology aircraft. One of the main advantages of using the logistic method, as opposed to a linear one, is that there are fewer requirements in terms of the assumptions in a logistic regression than for a linear regression (Cohen et al., 2003). In logistic regression, "the predictors do not have to be normally distributed, linearly related, or of equal variance within each group" (Tabachnick & Fidell, 2007:437). The logistic process was also used for its inherent robustness when dealing with dichotomous or ordinal dependent variables (Leech et al., 2005). Ho (2006) asserts that the logistic regression method is ideal for the analysis of perception instruments that yield dichotomous scores.

The model in this particular exploration was used primarily to uncover further important phenomena in the data set. The statistical analyses thus far had revealed some important outcomes; however, uncertainty remained about the relative predictive power of the independent variables. Some authors point out that one of the drawbacks in conducting a stepwise logistic regression is that it may model noise into the equation (Cohen *et al.*, 2003; Field, 2005; Goldstein & Wood, 1989). Such noise may result in further uncertainty in conclusions about the data set. Hence, it was important to discover the predictive power of variables from a regression analysis in order to add to the theoretical knowledge on the topic of interest.

Logistic regression was deemed suitable for this final examination of the research construct, because the dependent variable to be studied was a dichotomy and the independent variables were of varying types. The exploratory method involved an analysis, which began with a full or saturated model; variables were then eliminated from the model in an iterative process.

The aim in this stage of the study was to develop a model that could, to some extent, predict whether pilots would perceive the technologically advanced aircraft training climate as favourable or unfavourable, based hitherto on what is currently known about the latent structure of the dataset. Dummy variables of "1" or "0" were allocated to cases where the average perception on the three-factor model exceeded 5.0

(because of the right skewness in the final dataset) or not, as the case may be. This provided a dichotomous measure, *Favourability,* of a categorical outcome that indicated level of airline pilots' comfort in the advanced aircraft training climate in terms of their overall perception. A backward stepwise logistic method was used to determine the contribution each variable would make to the regression equation because "sequential logistic regression should be part of a cross-validation strategy to investigate the extent to which sample results may be more broadly generalized" (Tabachnick & Fidell, 2007:456). The equation for this exploratory method starts with all the selected independent variables first entered and then deleted after evaluation (Field, 2005; Ho, 2006). The following nine categorical or ordinal variables were included as predictors in the original model:

- The interaction effect between a pilot's level of flight experience in advanced aircraft with that of his or her level of computer literacy, where

  1 = Low Experience*Low Computer Literacy;

  2 = Low Experience*High Computer Literacy;

  3 = High Experience*Low Computer Literacy; and

  4 = High Experience*High Computer Literacy.

- A bivariate age grouping categorised in terms of respondents who indicated that they were either 40 years of age or younger and those who were over 40 years old, where

  0 = 40 and younger; and

  1 = 41 and older.

- Practical flight experience in advanced aircraft, where pilots with over 2 000 hours flight time on these aircraft were considered as having high experience, and where

  1 = Low Experience; and

  2 = High Experience.

- Route training, which indicated how much a pilot enjoyed training on the actual aircraft, where

  1 = Never Enjoy;

2 = Sometimes Enjoy; and

3 = Always Enjoy.

- Simulator training, which indicated how much the pilot enjoyed training experience in the flight simulator, where
  1 = Never Enjoy

  2 = Sometimes Enjoy; and

  3 = Always Enjoy.

- A variable which indicated the size of the carrier for which the respondent was employed, and where the size of the enterprise is represented as
  1 = Large;

  2 = Medium; and

  3 = Small.

- Pilot unionisation, which refers to the degree to which the pilot group was unified in terms of belonging to the ALPA-SA – the organisation was either unionised, or not unionised, where
  1 = Unionised; and

  2 = Not-unionised.

- The level of computer literacy, which indicated the extent to which the candidate felt that his or her computer skills were either poor or good, where
  1 = Low Literacy; and

  2 = High Literacy.

- A company status/position category, which divided the sample into two specific groupings, either captain or co-pilot, where
  0 = Co-pilot; and

  1 = Captain.

The backward stepwise regression analysis was completed after five steps. A model containing four predictors subsequently emerged. In addition, five of the aforementioned predictor variables were removed iteratively, namely pilot unionisation, age group, the size of carrier, the company status/position of the pilot and the pilot's actual perceived level of computer literacy. The four variables that successfully predicted a subject's perception of the *Favourability* associated with the advanced aircraft training climate were:

i) the interaction effect between a pilot's level of flight experience in advanced aircraft and his or her level of computer literacy:
ii) practical flight experience in advanced aircraft;
iii) preference regarding training in the flight simulator; and
iv) preference regarding route training in the actual aircraft.

Table 60 shows that the overall percentage of cases for which the dependent variable was correctly predicted by the model was 63.8%. The results also show that the model was correct in predicting perceptions of a *favourable* climate 100% of the time (that is, high positive predictive validity). However, the model could not successfully predict respondents' perceptions of an unfavourable training climate.

The effect size and practical significance of a regression analysis is generally provided for by examining the odds ratio and, less commonly, by computing the differences between Nagelkerke's pseudo R-Square values ($R^2\Delta$) (Cohen *et al.*, 2003). A cut-off value of $R^2\Delta = 0.02$ was used, as suggested by Schaap (2011), because the pseudo R^2 is not technically a goodness-of-fit index, and cannot explain the proportion of the variance. Therefore, the result based on computation of $R^2\Delta$ for this study was used with caution in interpreting practical significance

Nagelkerke's $R^2\Delta$ was computed by first calculating the value of the pseudo R^2 at the initial step and thereafter finding the difference at each subsequent step. The effect size of the model at each subsequent step was less than 0.02 and could therefore be regarded as not practically significant in terms of this criterion. Nonetheless, overall, the final model was regarded as efficient (see Table 60).

Table 60: Classification table and model summary

Classification		Predicted		
		Favourability of training climate		
Observed		Climate unfavourable	Climate favourable	Percentage correct
Favourable training climate	Climate unfavourable	0	83	(none correctly identified) 0.0
	Climate favourable	0	146	100.0
	Overall percentage			63.8
Model summary				
Step	-2 Log likelihood	Cox & Snell R^2	Nagelkerke's R^2	Nagelkerke's $R^2\Delta$
1	266.955	0.134	0.184	-
2	267.024	0.134	0.183	0.001
3	267.534	0.132	0.181	0.002
4	268.678	0.127	0.175	0.006
5	270.021	0.122	0.168	0.007
Step	Hosmer and Lemeshow test chi-square	Df	Sig. (2-tailed)	
1	13.884	8	0.085	
2	8.399	8	0.396	
3	13.720	8	0.089	
4	11.272	8	0.187	
5	2.365	7	0.937	

The efficiency of the resulting model is endorsed by the non-significance in the result of the Hosmer and Lemeshow test chi-square statistic in the final step (χ^2 [7, N=229] = 2.365, p = 0.937). These results confirmed that the final variables in the model predicted the observed data relatively effectively. According to Field (2005), the proportion of variance in the outcome variable associated with each of the predictor variables can be given by R^2. However, because the dependent variable in this regression model was dichotomous or categorical in nature, only approximations of R^2 were possible in SPSS Version 17.0; hence the choice of Nagelkerke's pseudo R^2 to gauge effect. Additionally, both the Cox and Snell $R^{2,}$ together with the Nagelkerke

R^2 values, were used to determine that the final model could account for approximately 12% to 17% of the variability in the criterion variable.

Table 60 shows that moderate changes occurred in the -2 log-likelihood values between the constant only model, and the first and last step, which is a good indication that the final model had improved predictive power. A nominal regression of the final four predictor variables in the model then produced a comparison in which the -2 log-likelihood values of the intercept only (95.338) and final model (65.456) indicated that the change in the amount of predictive power provided in the final solution was statistically significant [χ^2 (4) = 29.883, $p < 0.0001$].

Moreover, McFadden's p^2 [1-log likelihood (final)/log likelihood (constant)] = 0.313 was computed as an indication of a measure of the strength of association between the predictor variables and the model. According to Tabachnick and Fidell (2007:460), McFadden's p^2 is expected to be "lower" than the traditional R^2 as a measure of effect size, and values between 0.20 and 0.40 are considered highly satisfactory. In terms of McFadden's p^2 (as opposed to Nagelkerke's $R^2\Delta$), the analysis suggests that the size of the final logistic model is large and of practical importance.

Table 61: Final logistic regression model

Predictors in the equation (X_j)	B	S.E.	Wald Chi-square ($B^2/S.E.^2$)	Df	Sig.	Odds Ratio (e^B)	95% C.I. for odds ratio	
							Lower	Upper
Interaction effect	0.630	0.310	4.126	1	0.042	1.878	1.022	3.448
Advanced aircraft experience	-1.064	0.613	3.011	1	0.083	0.345	0.104	1.148
Enjoy route training	0.485	0.289	2.814	1	0.093	1.624	0.922	2.861
Enjoy simulator training	0.806	0.267	9.138	1	0.003	2.238	1.327	3.773
Constant	-2.603	0.912	8.142	1	0.004	0.074		

In the final model, the dependent variable, *Climate Favourability*, is on the logit scale. The computation results in Table 61 show that the probability that a respondent would perceive the advanced aircraft training climate as favourable can be given by the following three logistic regression equations (Tabachnick & Fidell, 2007:438):

- Logit = ln (p/1-p) = -2.603 + 0.63 * (interaction effect) – 1.064 * (advanced aircraft experience) + 0.485 * (route training) + 0.806 * (simulator training) …
[Equation 1]
- Prob (Favourable perception) = $(e^{-2.603 + 0.63 X_1 - 1.064 X_2 + 0.485 X_3 + 0.806 X_4})/(1+ e^{-2.603 + 0.63 X_1 - 1.064 X_2 + 0.485 X_3 + 0.806 X_4})$ …
[Equation 2]
- Log odds ratio = p/(1-p) … *[Equation 3]*

The model produced by the logistic regression is non-linear. In the current case, Table 61 shows that experience in an advanced aircraft and preference for route training did not improve the predictive efficiency of the model. The Wald chi-square statistics were non-significant for these two variables ($p > 0.05$), whereas the chi-square value for the interaction effect of advanced aircraft experience and computer literacy as well as for enjoying simulator training was significant at a 0.05 level. Nonetheless, Hosmer and Lemeshow (2000), also cited in Tabachnick and Fidell (2007:456), recommend that for "a criterion for inclusion of a variable that is less stringent than 0.05 …something in the range of 0.15 or 0.20 is more appropriate to ensure entry of variables with coefficients different from zero".

Thus, given that the other predictors remain in the model, it would appear that removing the interaction effect or information about preference for simulator training would result in significantly poorer predictive efficiency of the model (because these variables have greater significance). Overall, it can therefore be concluded that the predictive efficiency of a four-predictor model is not noticeably greater than that of a two-predictor model, which would then include only the variables *Interaction Effect* and *Preference for Simulator Training*.

The estimated logistic regression coefficient for the interaction effect between flight experience in advanced aircraft * level of computer literacy was 0.63 and the exponential of this value is $e^{0.63}=1.878$. This indicates that for a one-unit increase in the interaction effect, the odds in favour of a trainee’s perceiving the training climate as favourable are estimated to increase by a multiplicative factor of 1.878. In addition, $(p/1-p) = 1.878$ implies that knowledge of the interaction effect can improve the probability of correctly predicting that a pilot will perceive the advanced aircraft training climate as either favourable or unfavourable by 65%.

The estimated logistic regression coefficient for pilots’ preference for simulator training was 0.806 and the exponential for this value is $e^{0.806}=2.238$. This indicates that for a one-unit increase in a pilot’s positive preference for simulator training, the odds in favour of the trainee’s perceiving the training climate as favourable are estimated to increase by a multiplicative factor of 2.238. Furthermore, $(p/1-p)=2.238$ implies that having knowledge of a pilot’s positive preference for, or an enjoyment of simulator training, can improve the probability of correctly predicting whether he or she will perceive the advanced aircraft training climate as favourable by 69% (that is, $p = 0.69$).

To test the predictive power of the model, Equation 2 was then applied to the following extreme scenarios:

- The computation of the probability of a trainee’s perceiving the climate as favourable given that the pilot has a low advanced aircraft experience*low computer literacy combination, low advanced aircraft flight experience, and reports never enjoying route or simulator training:

  $$\text{Prob(Fav)}=(e^{-2.603 + 0.63 (1) - 1.064 (1) + 0.485(1)+ 0.806 (1)})/(1+ e^{-2.603 + 0.63 (1) - 1.064 (1) + 0.485(1)+ 0.806 (1)})$$

  $$=e^{-1.746}/1+ e^{-1.746}$$

  $$=0.174/1.174 = 0.148$$

  Therefore, the model predicts a very low probability (14.8%) that the candidate, given the aforementioned criterion, will perceive the advanced aircraft training climate as favourable.

- The computation of the probability of a favourable climate perception by a trainee who has a high advanced aircraft experience*high computer literacy combination, and high advanced aircraft flight experience, and who always enjoys route or simulator training:

  $\text{Prob(Fav)}=(e^{-2.603 + 0.63 (4) – 1.064 (2) + 0.485(3)+ 0.806 (3)})/(1+ e^{-2.603 + 0.63 (4) – 1.064 (2) + 0.485(3)+ 0.806 (3)})$

  $=e^{1.662}/1+ e^{1.662}$

  $=5.270/6.270 = 0.840$

  Therefore, the model predicts a very high probability (84%) that the candidate, given the aforementioned criterion, will perceive an advanced aircraft training climate as favourable.

The illustration of the usefulness of the logistic model as a predictor of climate favourability is demonstrated in the above example. Predictive efficiency thus appears to be very good for the two extreme cases.

The probability curves based on the derived probability formula are depicted in Figures 28 and 29. The plots were computed for the two most important independent variables found by the regression model:

- the *Interaction Effect*, where the variable is based on a combination of an airline pilot’s number of hours experience in advanced aircraft and perceived level of computer literacy; and
- a preference *for Simulator Training*, a variable that refers to the airline pilot trainee’s enjoyment of flight simulator training.

Figure 28: Probability plot for the interaction effect between experience and computer literacy

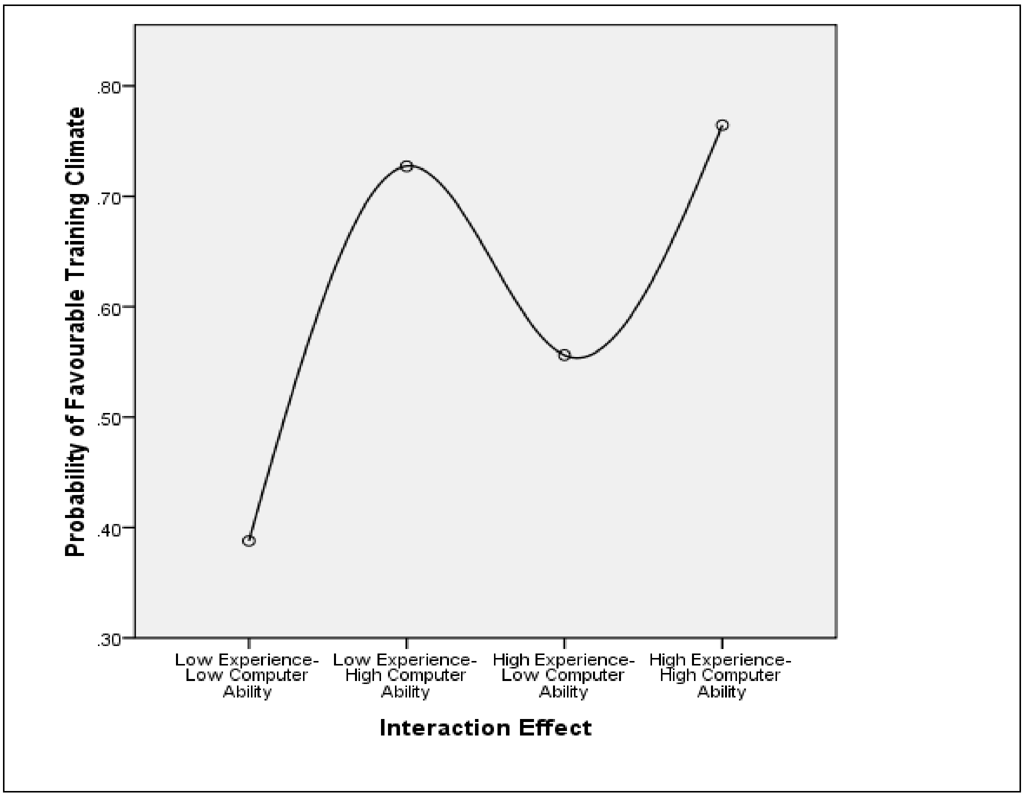


Figure 28 clearly shows the importance of a pilot's perceived computer literacy in the interaction effect with the candidate's advanced aircraft flight experience. The noticeable impact of this interaction on the probability curve is evident in the centre of the plot. The probability that a trainee will perceive the advanced aircraft training climate as favourable diminishes significantly when his or her levels of computer literacy drop, even when the person's levels of flight experience is higher (that is, the high probability of a favourable climate decreases between low experience*high computer literacy and high experience*low computer literacy).

The second important predictor in the regression model is plotted in Figure 29.

Figure 29: Probability plot for Preference for Simulator Training

The graph (Figure 29) depicts an almost linear relationship between a pilot trainee's preference or enjoyment of flight simulator training and the probability that the subject will perceive the training climate as either favourable or unfavourable. It appears from Figure 29 that understanding the levels of enjoyment a subject may experience with regard to simulator training has a significant impact on predictive power.

The confidence intervals of the odds ratios for all the remaining singular predictors for the final model were consistently positive. According to Field (2009), findings based on the results of reporting positive confidence interval values present a possibility to generalise a regression model to the broader population. However, the small values of Nagelkerke's $R^2\Delta$ in the current analysis imply that the practical significance of the model should be considered in the context of the shortcomings of the study. Moreover, the relatively high standard errors (SE) of each predictor variable point to the possibility that the model may be unstable due to interference or noise (Tabachnick & Fidell, 2007).

The high homogeneity of the sample frame in terms of their education, experience and training profoundly skewed the data that dominate in the final model, creating interference in the predictors. Nonetheless, the results and conclusions drawn from the regression analysis provide pertinent information for assessing the suitability of candidates for advanced aircraft training. This study supports the conclusion that knowledge of both the interaction effect of pilots' experience in advanced aircraft combined with a pilot's computer literacy or competence, and the knowledge of the levels of enjoyment a pilot derives from simulator training, are highly useful predictors of a trainee's perception of the advanced aircraft training climate. Appropriate interventions based on the understanding of this knowledge may in turn affect the level of success of the education methods and the interventions that are chosen, and thus the eventual success of training outcomes.

5.11 SUMMARY

The focus in this chapter was on interpreting and reporting the results of the phenomena associated with the latent factors of the theoretical construct. The results of a thorough exploration based on principal axis factoring and oblique rotation produced more factors than had been anticipated. Horn's parallel method, based on the algorithm provided by O'Connor, resulted in a three-factor final solution. It also eliminated the Heywood anomaly in the data. The latent behavioural scales were labelled *Organisational Professionalism*, *Intrinsic Motivation* and *Individual Control of Training Outcomes*. The first factor is related to both the macro (organisational) and intermediate (instructor-trainee) levels of analysis. The remaining two latent factors both entail variables related to the micro or individual level of analysis, suggesting that the trainee, in this case, is a focal point in the phenomena associated with training for operating advanced aircraft. The various analyses conducted suggest that the measurement construct under investigation has a relatively stable factor solution.

A reliability analysis of the final item cohort produced highly satisfactory Cronbach's coefficient alphas of between 0.70 and 0.95. It was also shown that the scale demonstrated very good discriminatory properties, given its ability to distinguish effectively between high and low scorers, suggesting excellent scale reliability and good overall development.

Exploration of the data set revealed that the scores were non-normal; therefore, subsequent analyses were based on using the appropriate family of non-parametric statistics. To establish a foundation for a further exploration of phenomena present in the data, a basic overall non-parametric comparative assessment was conducted, using the Kruskall-Wallis test statistic and supplemented for specific group comparisons by the Mann-Whitney test statistic.

For associational evaluation of the data, the non-parametric equivalent of Pearson's R, Kendall's tau-b, was deemed most appropriate for the exploration of correlations. The robustness of Kendall's tau-b assisted in offsetting the inherent problems associated with the non-normality and subsequent violations of assumptions in the current data set. The associational computation of conceptually significant variables allowed for a further statistically complex examination of the underlying phenomena present in the data. The results with regard to statistically significant correlations guided the researcher in the appropriate direction.

Subsequently, a multivariate analysis of variance was conducted on the ranked order of the scores on digital flight deck experience, a pilot's company status, the size of the carrier, pilots' levels of computer literacy and an interaction effect between experience in advanced aircraft and computer literacy. The results were then scrutinised and thoroughly discussed in a *post hoc* analysis.

In order to determine the predictability of the independent demographic data and to develop a structured explanatory model, a stepwise logistic regression analysis was carried out. The results produced a highly satisfactory model with four predictor variables. It was found that the derived logistic equation could satisfactorily compute, with reasonable probability (0.148 at the low end to 0.840 at the high end), whether a subject would hold an unfavourable or a favourable perception of the training climate.

Overall, the results indicate that, although the nature of the data was differentially skewed, the use of appropriate (more robust) statistics provided satisfactory answers and revealed distinct and important phenomena, thus making a significant contribution to the current theory base on the topic. Furthermore, the methodology followed exposed both a relatively stable underlying factorial structure and a prediction model that will be useful to airline organisations engaged in training advanced aircraft pilots.

One of the main findings from the data analyses is that the interaction between a pilot's level of advanced aircraft experience in combination with his or her computer literacy, competence and/or abilities, has a significant impact on the pilot's perception of the advanced aircraft training climate. In addition, it was found that the use of flight simulators plays an extremely important role in training perceptions. Further study linking synthetic flight training (simulation) to real-life aircraft training and operational behaviour is recommended in order to enhance the current theoretical knowledge base.

Against the background of the aforementioned research findings, a discussion of the results and recommendations follow in the final chapter.

CHAPTER SIX:

SUMMARY, DISCUSSIONS AND RECOMMENDATIONS

6.1 INTRODUCTION

This study has focused on the quantitative development of a valid and reliable instrument to measure airline pilots' perceptions of the training climate associated with advanced automated aircraft. The research was conducted using a positivist epistemology. Ultimately, the goal was to construct a questionnaire by operationalizing an unobserved hypothesised construct (perceptions of the advanced automated aircraft training climate) based on three levels of analysis (the microsphere, mesosphere and macrosphere) to conceptualise it. As a secondary purpose, the study sought to explore the relationships between demographic categories and phenomena relating to the data in terms of the latent sub-constructs that emerged from the empirical observation.

A preliminary literature review revealed that, although many airlines have initiated major purchases of advanced aircraft, very little research has thus far been undertaken to determine what the impact of the new technology would be on human behaviour in terms of aviation safety. The fact that human factor issues associated with advanced aircraft training are currently understudied provided the impetus for the present research. In particular, the topic of training pilots to operate technologically advanced aircraft has hitherto been neglected from a psychological and human behavioural perspective.

Global economic pressures have led aircraft manufacturers to turn to computer technology in the hope to create highly efficient and, ultimately more marketable products and services. The cost of this increased efficiency of aircraft in all respects, from surgically precise long-distance navigation to increased fuel economy, has led some observers to warn that human operators are being designed out of the flight deck and, more concerning, being designed out of the knowledge loop in general. The philosophy of current advanced aircraft manufacturers is that the less detail a pilot knows about the modern aircraft, the less likely they are to make mistakes. Overall, the statistics do tell a story of reduced human error, coupled with a reduced air accident rate. Detractors to increased aircraft automation however, continue to argue that the correlation between increased computerisation and reduced aircraft accidents rates, do not necessarily confirm causation, *per se*. These and other such issues in the industry, added to the interest in pursuing the present research study.

Computerisation and technology are a part of everyday life and are here to stay. This is particularly true for the commercial aviation sector. The implementation of technological advances in modern airliners has resulted in many improvements to flight safety, which has never before been possible. Paradoxically, increased computerisation has also introduced brand new human factor issues into the human-machine interface. These issues threaten to compromise the delicate fabric of aviation safety, and continue to profess the old adage that the human being will always be the weakest link in a complex system.

To mitigate adverse effects stemming from the complication arising from highly complex automation systems, research into training for advanced aircraft is increasingly being recognised as a way to reach a compromise between new computer-based aircraft and old human habits. This paradigm begins to challenge the problem at a root level. Because suitable training is a prerequisite for changing behaviour in an unusual workspace such as the flight deck, researchers in the field can add significant value to such specific flight safety initiatives.

The ability of an individual to navigate successfully the labyrinth of learning required to operate a highly advanced aircraft that employs complex automation depends on the person's perceptions of the training climate associated with these aircraft. This study was, in part, motivated by the realisation that only by understanding the psychological perspectives of trainees (which is a core individual level of analysis), whom are engaged in learning to fly advanced aircraft within an airline setting, is it possible to gauge the related phenomena associated with flight deck behaviour. In other words, this study tackled a root cause for pilots' flight deck behaviour (be it appropriate or inappropriate), namely, behaviour as a manifestation of training. The study concluded that, to a large extent, trainees' success in flight training depends on the way they perceive their training environment at three systemic levels (micro, meso and macro).

Neither safety issues nor the considerable financial implications associated with high training failure rates can be overemphasised; therefore it may be important for organisations to evolve training paradigms and methods by introducing appropriate interventions and policies before problems arise. The study focused on scale development, because accurate quantitative measurement of the climate can guide such interventions. Only by understanding empirically the phenomena associated with the psychological and in turn behavioural factors that affect the human-advanced machine interface, can any training intervention be successful in the long term.

As part of promoting an empirical ontological understanding of training for advanced aircraft, this study researched and designed an appropriate measurement instrument in terms of the positivist paradigm. Therefore, the quality of conclusions and recommendations may be regarded as relatively high, based on the stringent scientific method complied with. Data collected for this study can, and should be re-analysed, and findings corroborated or disputed by future researchers interested in the topic, or wanting a deeper understanding of aviation human factors related to technology.

This chapter contains a review of the results and a re-examination or discussion of some of the most crucial findings of this study. In order to completely understand recommendations and suggestions for future research; the study objectives, findings, the aims, assumptions and research methods that were dealt with in the first five chapters are revisited and revised for additional clarity. In addition, some of the implications for airline organisations and possible safety-enhancing recommendations are briefly discussed, based on only the main and significant findings in the study.

6.2 RESEARCH OBJECTIVES

In general, the purpose with this study was to measure the perceptions of airline pilots with regard to the advanced aircraft training climate and to compare and model these perceptions in terms of demographic variables. The study undertook to obtain sufficient empirical data from a representative cross-organisational sample of qualified airline pilots who were engaged in training to operate, or who had trained to operate, highly advanced commercial aircraft that employed very complex automation systems. The study specifically sought to measure the group's perceptions associated with the advanced aircraft training environment or climate, and to determine what constituted the latent structure of a hypothesised construct related to this climate. To follow up on the aforementioned discussion, the study also attempted to identify the underlying characteristic phenomena present in the data set by means of an intentionally thorough, statistical examination of the latent factors that were extracted from the data set. The main reason for this approach was based on an attempt to cover as much of the topic as feasibly possible so as to add to the body of knowledge with new information, and possibly assist future similar research.

The primary objective of the study was to obtain an empirical measurement of the hypothesised construct (see Figure 13) by conceptualising a multi-dimensional questionnaire in order to develop a valid and reliable scale. Secondarily, the research objective was also to explore the statistical nature of the demographic variables of the respondents in terms of the latent factors of the construct.

The following research objectives were generated to guide the study:

- to identify the organisational behaviour attributes applicable to the main research construct by critiquing and discussing the relevant literature, historical data, and current world trends on the topic of interest;
- to develop a hypothesised multivariate psychological systems model (based on empirically grounded theory) from which criteria for the construct *perceptions of the advanced automated aircraft training climate* can be identified and tested in a quantitative study involving a cross-organisational sample of airline pilots from South Africa;

- to generate a tentative pool of scale items based on the hypothetical systems construct;
- to validate the item pool statistically by quantifying the judgements gained from subject matter experts and using Lawshe's (1975) content validity ratio technique;
- to obtain sufficient empirical evidence to determine the nature of the latent factors of the hypothesised construct and to develop an understanding of their relationships to the surface attributes and to each other;
- to use statistical methods to develop a valid and reliable measurement instrument to assess airline pilots' perceptions of the training climate associated with advanced automated aircraft, where the following generic steps were used as a guide for the scale development (DeVellis, 2003; Netemeyer *et al.*, 2003; Pett *et al.*, 2003):
    - develop a hypothesised model of the main research construct;
    - generate appropriate items from the available theory;
    - operationalize the main construct by developing a scale (for instance, using results from an expert questionnaire); and
    - evaluate the robustness of the scale (appropriate statistics to determine validity and reliability);
- to explore the relationship between respondents' demographic variables and the latent factors of the construct;
- to determine what the theorised construct may be for future research (see Figure 60, for an example) based on the mathematical or statistical analyses of a dataset obtained through empirical methods; and
- to make recommendations to airline organisations and other interested parties based on the findings, which emerged only after a thorough or in-depth examination of the empirical dataset.

6.3 RESEARCH METHODOLOGY AND PROCEDURE

The primary objectives of the study were accomplished both from a theoretical perspective (in the form of the literature review) and in terms of practical assessment (the collection of empirical evidence). Defining the main research construct proved challenging, as a multitude of climate constructs were gleaned from the current body of knowledge. In developing the final training climatic construct however, three fundamental aspects were married, namely; first concepts, theories and constructs from education (learning), second, advanced aircraft technology and, third, climate theory. The “training climate” construct was by far the most complex definition to formalise for the study. It was found that a plethora of definitions, research and journal papers exist on the subject. Because most of the information in the literature on the topic deals with “training” and “climate” as separate constructs, determining what constitutes the applicable “training climate” proved challenging.

Adopting the positivist’s paradigm, a two-step multiple-method approach was used to validate a set of questionnaire items that operationalized the main research construct and ultimately provided the data to develop a reliable measurement scale. Therefore, two surveys were triangulated within a single ontology.

An expert survey was used to validate the formulation of the original 106 items of the hypothesised construct. Using a convenience sampling method, an expert group in the realm of interest was identified: 17 highly experienced airline flight instructors and university academics participated in the expert validation process. An instrument in the form of a structured questionnaire was developed and subsequently used to elicit judgements from the sample of experts. The content of the opinions from the experts was analysed in detail using the quantitative method developed by Lawshe (1975). The technique was successfully used to determine the extent of the overlap between the items formulated on the basis of the literature study and the domain of the hypothesised research construct.

The construct validation process resulted in a very rigorous elimination of redundant items. The panel of experts could agree on only 42 items of the 106 items initially presented for judgement which they believed adequately covered the topic of interest, and which could then be used in the next phase of scale development. This represented an endorsement rate of 39.62%. Although the endorsement of statements that could potentially define the main hypothesised construct was disappointing, the number of statements retained was sufficiently robust to continue with the research study. This was in accordance with the literature supporting the requirement for a minimum of 40 items in scale development of this nature. The very critical judgement received from the panel of experts resulted in a latent structure of the construct that may be considered very narrow in relative terms. Nonetheless, the final output, using 42 items, resulted in high communalities. Because the method was so stringent in nature, it becomes very difficult to dispute the point that the penultimate explanation of the variability in the construct under study was not derived from a high quality measurement scale.

The final 42-item cohort was packaged in a structured questionnaire and distributed both electronically and manually to the population of interest. However, the web-based electronic questionnaire was used as the main source of empirical data. Both airline management and pilots' unions were extremely helpful in making this study a success by providing the required assistance to elicit participation in the survey. A total of 229 useable returns formed the basis of the final exploration, examination and analyses of the quantitative section of the research.

The data was audited and presented in an appropriate format for analysis. Two statistical packages, namely Statistica Version 7 and SPSS Version 17, were used for the data analysis. The analytical process began with a basic factor analytic exploration, namely a principle axis factoring. This analysis produced a stable three-factor solution. The latent factors that explained most of the variability in the construct were labelled *Organisational Professionalism*, *Intrinsic Motivation*, and *Individual Control of Training Outcomes*. Firstly, the factor called *Organisational Professionalism* was formulated based on items describing both the macro level of analysis (the organisation or airline) and the meso level of analysis (the group or instructor-trainee combination). Secondly, the final two factors (*Intrinsic Motivation* and *Individual Control of Training Outcomes*) were formulated using items that described only an individual (or trainee) level of analysis.

Although the original hypothesised model of the research construct was based on three levels of analysis (the macro, meso, micro levels), interestingly, the final latent structure combined two levels (the macro and meso levels) into one (Factor 1) and split the core level (the micro level) into two factors (Factor 2 and Factor 3), clearly shown in Figure 30. This structure suggests the important role played by the individual in a systemic and dynamic environment, thereby conforming to the seminal premise of organisational behaviour, which starts at the individual level. In other words, actions or behaviour at the micro level permeate the other two levels, thus affecting the system holistically.

6.4 MAIN COMPONENTS OF THE HYPOTHESISED RESEARCH CONSTRUCT

A principal axis factor analysis with an oblique (promax, Kappa-4) rotation was conducted on the data to determine its underlying or latent structure. The main construct is finalised and theorised in Figure 30. In addition, the following main components emerged from this analysis (also see Appendix E):

- Factor 1 – *Organisational Professionalism*:

  This factor essentially relates to the organisational, managerial, bureaucratic and hierarchical aspects of the training climate. Items from both the macro or organisational level (that is, the airline) and intermediate or group level (that is, the instructor-trainees collective) dimensions loaded substantively onto Factor 1 (see Figure 30). Essentially, the factor expresses a component of the theoretical construct relating to the efficiency, effectiveness and professionalism of both the company and its flight instructors. The elements that dominate this factor include organisational co-ordination, structure, trainee support, rules, regulations, sufficient learner feedback and guidance. In terms of the main theorised construct however (see Figure 13, Figure 30 and Table 62), this factor included the domains; Structure (Str), Training Standards (Std), Training Culture (Cu), Training Planning (TPla), Knowledge Environment (En), Communication (Com), Training Conflict (Co), Training Policy (TPol), Intergroup Training Behaviour (InGB), and Power (Pr).

- Factor 2 – *Intrinsic Motivation*:

  This factor contains elements related to the individual or micro level of analysis, the trainee (see Figure 30). The factor predominantly reflects the individual trainee's ability and eagerness to learn and understand complex concepts relating to an advanced aircraft. Learning the aspects of a complex technology is viewed as a structured and iterative quantitative increase in knowledge. The fundamental aspects of this factor relate to an individual's learning approach, preparedness, and willingness to participate and co-operate in training to gain a knowledgeable and working understanding of the advanced aircraft that the person is being trained to fly. In terms of the main theorised construct however (see Figure 13, Figure 30 and Table 62), this factor included the domains; Motivation to Train (Mo), Personality (Per), Training Decision Making (Dm), Training Stress (Sts), and Learning for Technology (Le).

- Factor 3 – *Individual Control of Training Outcomes*:
  The third factor represents the trainee's own perceived level of control regarding their training to operate an advanced automated aircraft, and is measured at an individual level of analysis. Four items found at the micro or individual level (the trainee) loaded meaningfully onto this factor (see Appendix E). The principal elements of Factor 3 relate to the levels of perceived comfort experienced by trainees during training, their belief in their ability to control the outcome(s) of a training session, their capacity to control their levels of stress to perform well, and, ultimately their grasp of the full information load that they are required to cope with in their training. In terms of the main theorised construct however (see Figure 13, Figure 30 and Table 62), this factor included the domains Training Stress (Sts), and Training Decision Making (Dm).

The results of an exploratory factor analysis and item discriminant computation suggest that the final scale produced from the Advanced Aircraft Training Climate Questionnaire (AATC-Q) developed in this study has acceptable psychometric properties and can be confidently used to assess pilots' attitudes or perceptions regarding the advanced automated aircraft training climate for airlines based in South Africa (Appendix C).

It is suggested that airlines and interested parties use the measurement scale to

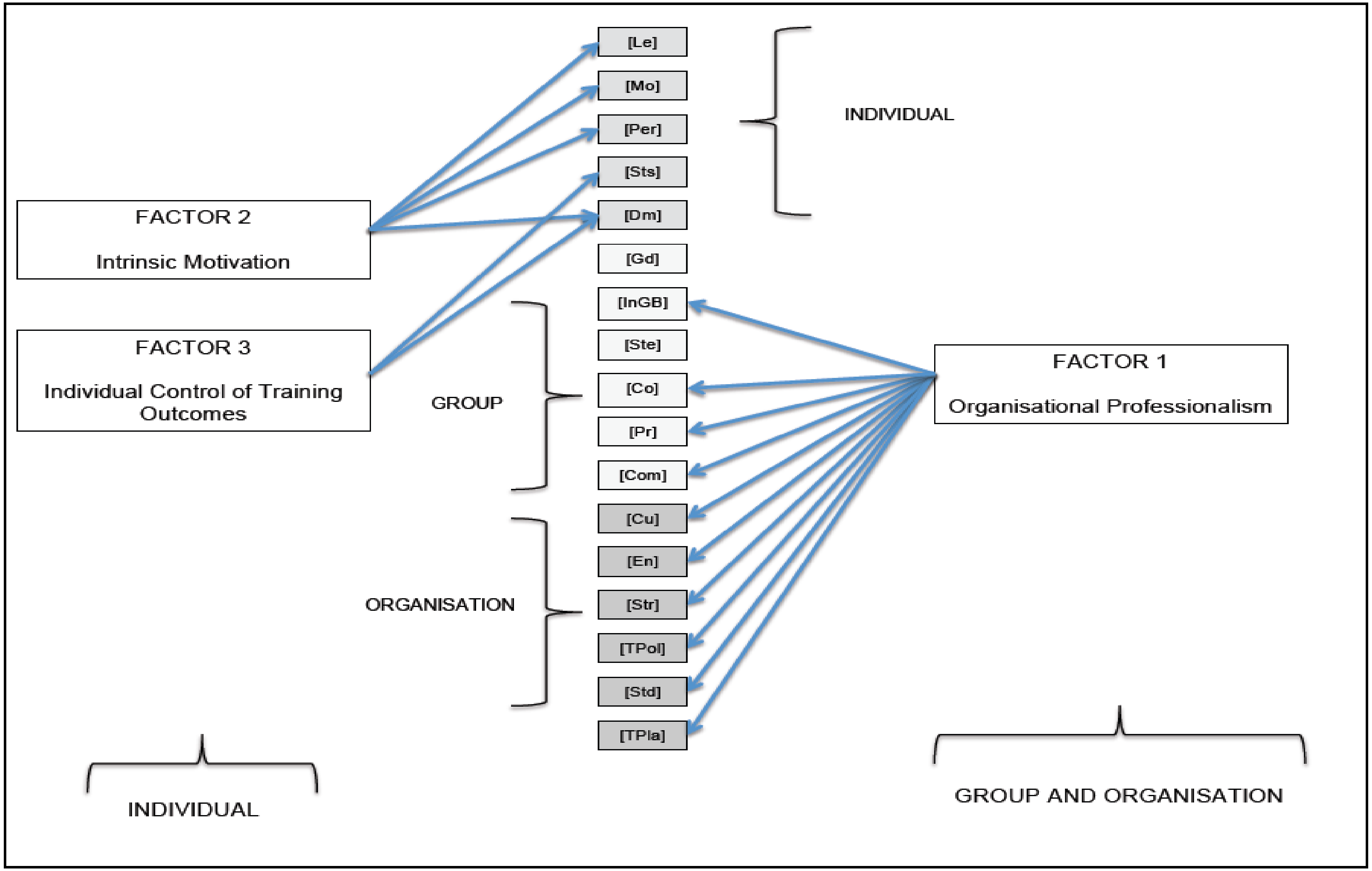
- sensitise airline pilots to their own perceptions with regard to training for technologically advanced aircraft employing highly complex automation systems and how these attitudes can assist or hamper transition training and its long-term success or failure;
- promote communication at three levels of analysis (the individual, group and organisation) in order to improve systemic understanding and thinking, which may ultimately enhance safer and effective flight deck behaviour;
- enhance overall policy changes, which can create an effective and sustainable learning environment for those learning to fly advanced aircraft; and
- assist airline organisations to develop strategies which will improve both aspects preceding training (such as recruitment) and the final training outcome success for advanced aircraft pilots.

The validity and reliability (Cronbach's $\alpha > 0.75$) of the derived behavioural scales (Organisational Professionalism, Intrinsic Motivation, Individual Control of Training Outcomes) were shown to exceed the recommended criteria. These results allowed the study to progress to the next stage, where analyses based on comparative and associational designs were conducted. Finally, more advanced mathematical modelling, based on a stepwise logistic regression computation, was conducted to enhance understanding of phenomena noted in the data, and to interpret the main and secondary findings.

Table 62: Relationship between construct domains and derived scales

Item	Scale Statement	Construct Domain
Q29	Training on this aircraft is well organised.	Structure (Str)
Q27	Training on this aircraft is professional.	Training Standards (Std)
Q23	My company's training produces world class pilots.	Training Culture (Cu)
Q24	Training at my airline is in line with company goals.	Training Planning (TPla)
Q38	The airline is very supportive of its pilots' learning requirements for this aircraft.	Training Culture (Cu)
Q34	There is sufficient training guidance from the company.	Knowledge Environment (En)
Q28	Management follows the rules and regulations appropriately.	Training Standards (Std)
Q39	My company's culture supports training for new technology aircraft.	Training Culture (Cu)
Q30	I understand what the company expects of me when training.	Communication (Com)
Q26	My company has talented people in training.	Knowledge Environment (En)
Q33	If I had to experience a problem in training, it's easy for me to appeal.	Training Conflict (Co)
Q25	I know what my company's training goals are.	Knowledge Environment (En)
Q31	Training at my airline produces safe pilots.	Training Policy (TPol)
Q40	There is sufficient feedback about my training on this aircraft.	Communication (Com)
Q42	My company uses only current training material.	Training Standards (Std)
Q41	Training is in line with civil aviation regulations.	Training Standards (Std)
Q32	The airline gives its pilots an appropriate amount of preparation work for training.	Training Planning (TPla)
Q45	My instructor is willing to listen.	Intergroup Training Behaviour (InGB)
Q50	Pilots are in direct control of the training outcome.	Power (Pr)
Q36	I'm given sufficient time to prepare for training on this aircraft.	Training Planning (TPla)
Q61	It's a good idea to know more than what is required.	Motivation to Train (Mo)
Q52	I try never to be late for a training session.	Personality (Per)
Q53	I co-operate when training in a simulator.	Personality (Per)
Q62	I aim to gain a deeper understanding of this aircraft.	Training Decision Making (Dm)
Q51	Preparation improves performance.	Training Stress (Sts)
Q60	I read to understand so as to gain a deeper understanding of this aircraft's systems.	Learning for Technology (Le)
Q55	I have a positive relationship with my colleagues.	Personality (Per)
Q44	I operate well as a crew member in the simulator.	Training Decision Making (Dm)
Q58	I enjoy studying the technical aspects of the aircraft.	Motivation to Train (Mo)
Q63	I'm comfortable undergoing training for this aircraft.	Training Stress (Sts)
Q57	I'm in control of the outcome of a training session.	Training Decision Making (Dm)
Q64	I can control my anxiety so as to perform well in training.	Training Stress (Sts)
Q49	The instructors on this aircraft don't overload us with information.	Training Stress (Sts)

Figure 30: Final theorised construct based on the empirical dataset



6.5 MANAGERIAL IMPLICATIONS OF THE RELATIONSHIPS BETWEEN DEMOGRAPHICAL AND OUTCOME VARIABLES

The study was conducted strictly within the quantitative, positivist paradigm. Therefore, the findings are based on empirical evidence and the conclusions are statistically supported by a p-value and associated effect size wherever necessary. Although the p-values of many of the findings were significant, the effect size in terms of Cohen's (1988) criterion was found to be relatively mediocre, possibly because the sampling frame was somewhat skewed and relatively small. Nonetheless, robust non-parametric statistical methods were adopted in the data examination to support and validate the conclusions that follow. The final analyses were thus regarded as useful and informative from a managerial and research perspective.

The study has enabled a deeper understanding of the underlying phenomena present within the current dataset, thereby also fulfilling the secondary research objective. Significant theoretical contributions to the current understanding of the topic were provided on the basis of following an in-depth analytical design. For instance, the Kruskal-Wallis test was conducted to compare multivariate data. In addition, where significant differences were found, *post hoc* non-parametric Mann-Whitney tests were used to determine the statistical significance of the actual difference between the highest and lowest ranked categories. Where the order of the independent categories was thought to be meaningful, the Jonckheere-Terpstra (J-T) test was used to assess the possibility of data trends. Furthermore, to assess the relationship between or the extent to which variables may be related and to determine their magnitude and subsequent direction, a correlational analysis was conducted. Kendall's tau-b was used for this measure of association. In addition, more advanced statistics were used to gain a deeper understanding of the dataset. For example, to examine the main and interactional effects of partially independent categorical variables on multiple dependent variables, a non-parametric MANOVA was computed (Pillai-Bartlett trace). Finally, a predictive model of the dataset was developed, based on a stepwise logistic regression. The main theorised construct could therefore be regarded as having attained high content, construct, and predictive validity.

The discussion below is based on the results of statistical analyses. The section is divided according to important demographic findings.

- ***Flight deck position***

  Although airline pilots can hold various positions in the organisation; to make further calculations easier, the pilots were initially captured as predominantly belonging to one of two main categories – either captain (commander of the aircraft) or co-pilot (second-in-command on the aircraft). These two main categories could then be sub-divided further into pilots who operated mainly long-range flights (flights more than six hours long) and those who operated only short-range flights. This sub-division provided an opportunity to analyse four significant demographic classes of airline pilots found in the dataset.

  A non-parametric comparison of the subgroups for the pilots' flight deck position showed that only the behavioural scale *Individual Control of Training Outcomes* was statistically ($p < 0.05$) affected by one of the four flight deck positions (captain long-range, captain short-range, co-pilot long-range, co-pilot short-range). The *post hoc* Mann-Whitney test showed that the short-range captains' scores were statistically higher than those of their long-range counterparts ($p = 0.001$). One possible reason for this difference may be based on the fact that short-range pilots have the opportunity to fly more sectors than long-range pilots do, which implies increased exposure to the technology for the short-range pilots. These findings corroborate an earlier study (Naidoo, 2008), which clearly showed that long-range captains (commanders) had a statistically significant lower perception of advanced aircraft automation in general, when compared to captains who operated the short-range advanced aircraft. However, such a phenomenon may be temporary, and isolated to the specific sample frame. Over time, short-range captains would become the new long-range captains, as they move up the seniority ranks. It can therefore be concluded that the results obtained from the present dataset, is due in part to the fact that the current generation of long-range captains within the present sample frame have high experience levels in analogue type aircraft. Experience levels in traditionally analogue type aircraft are expected to dissolve over time as airlines begin acquiring more efficient advanced digital aircraft. It would be a recommendation that the theorised construct developed for this study be re-examined in years to come, when it is expected that very few pilots would have any exposure to traditional analogue type aircraft.

In terms of the present study's findings, low perceptions of individual control of training outcomes may be enhanced if a pilot has opportunities to operate the aircraft in question more frequently. This is in accordance with Fishbein and Ajzen's (2001) proposal that familiarity or frequent exposure can enhance perception positively. It is therefore recommended that long-range captains who are not very experienced on particular types of aircraft or who have very little advanced aircraft experience, be given flight opportunities on shorter sectors. Possibly, these pilots should be rostered (planned/scheduled) to fly more often on domestic and regional flight sectors during the early phases of operation after their qualification on an aircraft. This will boost their personal confidence in operating the new advanced aircraft through exposure, and significantly improve their perceptions of control, resulting in safer and more competent pilots. In the long-term however, such intervention would not be necessary as all new long-range captains are then sourced from the short-range advanced aircraft pool.

- ***Size of the organisation***

  In order to compare different sizes of airlines found within the dataset, three categories were created to make an initial computation easier. Airline organisations were classified into one of three main groupings, depending on the number of aircraft the organisation operates and the number of employees working for the enterprise. Pitfield, *et al.,* (2010) describe carrier sizes in terms of the number of aircraft carrying more than 99 passengers in an operating fleet. The size of the organisation used for the present study was then categorised as small, medium or large (defined in Section 4.11). It was found that two behavioural scales were significantly affected by the size of the carrier demographic, namely *Organisational Professionalism* and *Intrinsic Motivation* ($p < 0.05$). However, *post hoc* testing showed that only perceptions of the behavioural scale *Organisational Professionalism* were significantly affected by whether a carrier was regarded as a relatively larger operator or as a relatively smaller operator. It was noteworthy that the pilots employed by the operators classified as “larger”, scored statistically significantly higher (with a medium effect size) on the *Organisational Professionalism* behavioural scale.

A possible reason for this observed phenomenon may be, that when employees find it easier to familiarise themselves with management and instructors, as in the case of smaller companies, they tend to have lower perceptions of the level of professionalism offered. In addition, larger companies also have access to more budget and therefore greater resources, creating a sense of professionalism. To maintain control over a large corporation it is possible that complex levels of bureaucracy evolve within the enterprise in a false hope in stemming resource haemorrhaging (Drucker, 1946), these bureaucracies may in turn add to a false perception of organisational professionalism for the individual. Alternatively, and more likely; when one compares the larger airline organisations in South Africa (South African Airways and British Airways Comair) to the smaller operators, it is noticed that the two larger companies utilise and maintain their own flight simulator centres. Pilots employed at the larger operators therefore undergo flight simulator training twice a year, as opposed to the smaller operators’ pilots who undergo flight simulator training only once a year. This difference may add to the higher levels of organisational professionalism perceived by the trainees of the larger operators.

It is recommended that the flight instructors and management in charge of training structure and policy at carriers with low scores on the *Organisational Professionalism* scale make an effort to communicate with trainees regularly and in a more structured, or formalised way. Training centres should be well equipped with the resources needed to enhance learning (such as training footprints, programmes, timetables, audio and visual media, and full time clerical staff). In addition, smaller airline operators of advanced aircraft should endeavour to send their pilots for flight simulator training at least twice a year (every 6 months) as opposed to the mandatory once a year requirement currently practiced.

Clearing up ambiguity and misunderstanding early on in a candidate's training experience by opening up communication channels, can also influence trainees' perceptions of organisational professionalism, and ultimately catalyse successful outcomes, not only for the individual, but also for the organisation systemically.

- ***Computer literacy***

It was important to understand and examine the effects of participants' perceived levels of computer literacy on their perceptions of the advanced aircraft training climate, because computer science is regarded as the cornerstone of modern technology.

A Jonckheere-Terpstra test suggested that the order of computer literacy – poor, average, above average or excellent – was statistically significant. Only the two behavioural scales at the individual level of analysis were affected by this demographic: *Intrinsic Motivation* and *Individual Control of Training Outcomes* (J-T[4] = 10178.50, 10888.0; Std. J-T[4] = 2.948, 4.272; $p < 0.01$). These results provided evidence which attests to the phenomenon that, as pilots' perceived levels of computer literacy increase or improve, so too does their motivation to learn about new technology aircraft and personal feelings about their ability to control outcomes related to training for advanced aircraft. The third latent construct, *Individual Control of Training Outcomes*, was significantly correlated with pilots' perceptions of their computer literacy (τ_S = 0.216, $p < 0.001$, small effect). This result indicated that, as pilots begin to believe that their levels of computer literacy improve, so too do their perceptions of their ability to be in control of, and take charge of their learning to operate advanced aircraft types.

It is recommended that airlines recruiting pilots to operate technologically advanced aircraft, should possibly consider a candidate's computer science background at either the secondary school or tertiary education levels. Candidates with such computer science backgrounds are likely to adjust more easily to the technical aspects of the advanced aircraft. It is also recommended that recruitment specialists at airlines be more cognisant of technologically averse candidates when hiring for the advanced aircraft fleets. Such candidates are likely to struggle through their conversions onto advanced aircraft employing complex automation. Subsequently, technologically averse pilots may become a safety issue (from a human factor perspective) on the actual aircraft engaged in line operations.

- ***Effect of Digital flight experience*Computer literacy interaction on perceptions***

Perhaps one of the more important findings in this study was that there was a significant influence from the interacting effect between subjects' levels of experience in advanced aircraft on the one hand and their perceived levels of computer literacy on the other, as opposed to the influence of the variables "experience" and "computer literacy" in isolation. It was noteworthy that the behavioural scale *Individual Control of Training Outcomes* was significantly affected by this interaction and that this in turn had an impact on respondents' perceptions of the advanced aircraft training climate in terms of the logistic sigmoid or s-curve.

After an examination of the data, it was found that the mean rank scores on the *Individual Control of Training Outcomes* (Factor 3) behavioural scale were significantly different for pilots who reported low experience in advanced aircraft, combined with a high level of computer literacy (Mean Rank = 82.48), when compared to the scores of pilots who reported high experience in advanced aircraft, combined with a low level of computer literacy (Mean Rank = 58.72). The effect of this difference in scores was regarded as medium ($U=1585.5$, $p < 0.001$; $r = 0.30$).

This finding reveals that, even though a subject may have very little experience on advanced aircraft, if the person has some computer science background (possibly at secondary school or university level), the candidate's lack of experience can be offset by high computer literacy levels, which in turn increase or improve personal control in training success. The ability to grasp abstract computer science concepts at a secondary school or tertiary education level (computer literate) is a definite indication of a candidate's ability to grasp technical and abstract advanced aircraft issues. As aircraft become more technologically advanced, such skills, knowledge and related attitude may become increasingly important in new recruits, because it would be a difficult if not an impossible task for any airline training section to cover every aspect or scenario associated with the technicalities of a new aircraft. Therefore, high levels of aircraft experience would not guarantee ease of training for a trainee who is computer illiterate. The inherent motivation of a candidate, as a spin-off from the interaction effect, can make the difference between highly knowledgeable and less knowledgeable pilots. An individual effort and personal responsibility of ground school studies are vital for a trainee of an advanced aircraft to succeed.

This conclusion and recommendation is regarded as important for airlines, particularly when recruiting new hire pilots for advanced aircraft training. Ensuring that candidates with the optimal combination of flying experience and computer literacy are employed may promote successful training outcomes. The effect of the interaction between levels of experience and perceived computer literacy was scrutinised in depth in Sections 5.9.4 and 5.10, using both multivariate techniques and logistic regression.

- ***Pilots' initial training***

  Traditionally, airlines have endeavoured to recruit military-trained pilots. The logic behind this method was that it could ensure safety within flight operations from a highly structured environment based on rules, regulations and standardised operating procedures. Therefore, it was hoped that military trained recruits would have developed advanced regimental skills from exposure to the military, in turn becoming ideal candidates for employment within such a highly structured organisation. The results from this study show that there may have been some merit in this preference, as a statistically significant difference was found between the perception scores of those pilots who had been trained in the military (in a highly structured, regimented environment [Mean Rank = 127.39]) and those of pilots who had no military background (Mean Rank = 106.57) with regard to their views of *Organisational Professionalism* (U =

4828.5, $p < 0.05$). However, these results should also be taken in context, in that the majority of military-trained airline pilots are also employed at the larger carriers (South African Airways and British Airways Comair).

Nonetheless, a regimented (that is, military) or otherwise highly structured initial flight training background can influence the overall perceptions of the advanced aircraft training climate for subjects who are making the transition to advanced aircraft. Pilots who are training for these types of aircraft may be required to adapt their knowledge of, or past experiences with, a structured learning environment to facilitate their acquisition of technically complex information. On the basis of this finding, it is recommended that airlines ensure that new candidates are provided with precise or detailed training plans prior to commencing their training for advanced aircraft. Such a structured approach to training for advanced aircraft can enhance organisational perceptions of the learning environment and thus result in more successful training outcomes.

- ***Predictive model***

To obtain a predictive model based on the empirical data, a backward stepwise logistic regression analysis was completed. After five steps, a model containing four predictors subsequently emerged. The process resulted in the identification of four demographic variables that successfully predicted a subject's perception of the *favourability* associated with the advanced aircraft training climate. The overall percentage of cases for which the dependent variable was correctly predicted by the model was 63.8%.

In the final model, the dependent variable *positive climate favourability* was calculated on the logit scale. The computation results then showed that the probability that a respondent would perceive the advanced aircraft training climate as favourable could be expressed with the following logistic regression equation:

$$\text{Logit} = \ln (p/1-p) = -2.603 + 0.63 * (\text{interaction effect}) - 1.064 * (\text{advanced aircraft experience}) + 0.485 * (\text{route training}) + 0.806 * (\text{simulator training}).$$

The logistic equation was an important result from the data exploration, as it provided evidence that, if an airline organisation knows a candidate's score on the interaction effect between the candidate's level of experience and his or her perceived level of computer literacy, his or her actual experience in advanced aircraft and his or her preferences for route and simulator training, the organisation can effectively predict the state of the training climate (level of favourability).

Alternatively, it can be deduced that the advanced automated aircraft training climate comprises the components, computer literacy*advanced aircraft experience interaction, route training preference, and simulator training preference. These four categories of demographics have emerged as important variables within the dataset. Hence, it is recommended that airlines make an effort to gain an understanding of these variables for their candidates when assessing the organisation's advanced aircraft training climate. Addressing any shortcomings associated with these demographic variables will enhance training plans, policies and structures, thereby strengthening the probability of a successful training outcome.

➢ ***Findings at the organisational behaviour levels of analysis***

At both a macro level or organisational level, and an intermediate or group level of analysis, an examination of the results suggests that it is imperative that management and flight instructors provide trainees with sufficient feedback and timeous learning plans. Therefore, the results highlight the point that communication can play a vital role in the advanced aircraft training climate. Effective communication based on effective feedback loops appears to be essential for creating a sustainable and favourable advanced aircraft training climate. Management should therefore explore various communication methods, such as electronic versus paper-based communication. Conversely, sufficient feedback should be encouraged from learners so that those who have the authority and power to do so have an opportunity to implement timeous changes or to enhance the training policy, procedure and structures systemically.

At a micro level or individual level of analysis, the data clearly shows that the latent structure of the main construct under examination consisted of two factors of importance based on the trainee. This provides sufficient evidence that the trainee is the cornerstone of the training initiative. Every individual pilot is therefore responsible for both the control over, and final success, of an advanced aircraft training event. Management and flight instructors should provide sufficient support in the form of learning and study skill-sets to trainees who appear to be having problems. Instructors of advanced aircraft should therefore be extremely good empathisers with an acute ability to understand others' stressors.

It is more than likely that candidates who are undergoing advanced aircraft transition training who are not technologically averse, and who end up being unsuccessful; simply do not have, or are not aware of, the tools available to them to facilitate their learning. In these cases, one-on-one instructor-trainee input is required in order to strengthen the micro-level and thus enhance the training climate as perceived by that individual.

6.6 LIMITATIONS OF THE STUDY

In general, the results suggest that the behavioural measures of *Organisational Professionalism*, *Intrinsic Motivation* and *Individual Control of Training Outcomes* are sufficiently reliable and valid to capture the perceived training climate associated with advanced aircraft that employ highly complex automation. Some additional elements that influence overall perceptions may relate to the type of organisation and the leadership style associated with it at both the managerial and instructional levels. Researchers should therefore endeavour to select appropriate measures that incorporate elements that are relevant to specific contexts, and be cautious in generalising the results of the AATC-Q across international contexts. A cross-national comparison and validation using this measurement instrument could solve some of the problems related to the context-sensitivity of the three scales. In addition, any additional examination of phenomena associated with demographics, requires a thorough definition of demographic categories.

One limitation in this study arises from the use of a highly comprehensive and structured questionnaire in both the initial construct validation and the large sample survey. Each questionnaire also consisted of a detailed preamble, requiring a specific level of understanding by the participant. Although the final questionnaire eventually consisted of fewer than a hundred questions, it still took approximately 30 minutes to complete. This could have influenced respondents' willingness to participate.

Another limitation concerned the interpretation of the questionnaire, since some candidates may have felt that the survey was connected with the management of their company. At some organisations, pockets of pilots were suspicious of management's intentions and were therefore, understandably very reluctant to participate in the study (even though such suspicion was unfounded). These splinter groups within the population may have caused an increase in response-based bias (where participants who completed the survey were more likely to answer questions in a particular way in an effort to please the researchers), and reduced the overall response-rate.

Although items to develop the scale were thoroughly researched and were based on previous studies of a similar nature in psychology, better items could arguably be selected, validated and found reliable in such measurement construction. Retention of several unidentifiable inferior items may be the reason why the distribution of scores of the factors was not normal and did not yield significant results. This is also reflected in the low practical significance values (effect sizes) found throughout the study.

6.7 RECOMMENDATIONS FOR FUTURE RESEARCH

The following recommendations are made on the basis of the findings:

- Perhaps the more disappointing aspect of this research project was the relatively low level of endorsement of items presented to the subject matter experts for content validation. Future research may be needed to redefine the operational nature of hypothesised construct. The universe of possible content that may operationalize the research construct is limitless, based on the depth and breadth of literature reviewed, therefore better item statements may be used for construct content validation. Nonetheless, the final operationalization of the hypothesised construct and empirical data used in the present research was adequate in constructing a main theoretical construct, for understanding the related phenomena, developing a valid scale and predictive model, and in turn adding new knowledge to the current understanding of training for advanced automated aircraft.

- Out of the over thirty demographic variables used in this study, only a few emerged as significantly and practically relevant to the differences between the various categories or groups regarding perceptions of the advanced aircraft training climate, these were:
  - a candidate's score on the interaction effect between his or her experience level and perceived level of computer literacy,
  - a candidate's actual experience in advanced aircraft,
  - a pilot's preference for route training,
  - a pilot's preference for simulator training,
  - size of the carrier,
  - flight deck position,
  - candidate's level of computer literacy,
  - pilot's initial training.

  Future studies could explore the effect of variables not yet considered in this study, such as a trainee's preferred learning style.

- Behaviour on the flight deck is still influenced by the ability developed by a pilot's Crew Resource Management (CRM) skill. All the airline pilots in the sample frame indicated that they had undertaken a formal CRM course. This may or may not have had an impact on the perception scales developed; however, the relevant literature presents sufficient evidence to suggest that CRM plays a role in enhancing flight safety. It is therefore recommended that future research also focus on the link between CRM and perceptions of the advanced aircraft training climate. Perhaps a new generation of CRM is required for technologically advanced aircraft. The management of resources, not only at a human level, but also at a machine (computer) level, is now a requirement. Present-day CRM targets the human-human interface, but so far, very few tools are provided to pilots to deal with the human-advanced machine interface.

- It would be of particular interest to validate the AATC-Q for internationally operated airlines based in other countries. Findings in this regard could provide the data necessary to produce a generic scale that takes nationality variables into consideration and could therefore also produce a different latent structure of the construct under investigation.

- Future research could use structural equation modelling methods to determine the validity of the three-factor model of the AATC-Q. Exploratory factor analysis, as used in this study, has its limitations in determining structural validity. Therefore, a structural equation modelling method may enable a researcher to postulate relationships between the observed measures and the latent factors *a priori*. The *a priori* relationship between the latent structure and observed variables should then be evaluated statistically to determine the goodness-of-fit with the empirical evidence. The findings could refute or verify the quality of the scale developed in this thesis and therefore provide more information to researchers in the field of aviation psychology, potentially enhancing flight safety from a human factors perspective.

- Additional research is definitely required with regard to the interface between human factors and technologically superior aircraft, both in commercial and general aviation. Longitudinal studies could provide valuable input and add to the current body of knowledge in aircraft automation issues. These studies could also be used to make adjustments to, or complete changes in, airline training philosophies.

6.8 CONCLUDING REMARKS

The application of advanced statistical procedures in this study was made possible by the various excellent commercially available computer software packages. This has allowed the thesis to examine thoroughly the validity of the multidimensionality of the hypothesised construct related to the perceptions of the climate associated with advanced aircraft training. Studies using such procedures and the results from the present study should be considered as only the beginning of the process of unravelling the complex aviation human factor issues associated with modern technologically advanced aircraft using precise measurement tools. The research has provided a logical positivist understanding of training for advanced aircraft and related phenomena. It is hoped that airline organisations, aviation psychologists and other interested parties will use the information set out in this thesis to enhance flight safety systemically – from initial recruitment to the final operational behaviour of the pilot on the flight deck. The findings of, and discussion in this thesis confirm that we have only scratched the surface of the human-advanced aircraft dynamic. It is likely that serious incidents and aircraft accidents associated with the interface between technology and human beings will continue to occur as systems become more and more complex.

Education and training in the appropriate use of such technology are the most logical ways available to airlines to combat the potentially adverse impact of human behaviour on the flight deck, reduce error and in turn reduce accident/incident rates. It is therefore vital that the industry (from regulators, to schools and associations, to enterprises) make a concerted effort to address the ways in which aviators in general, and airline pilots in particular, interact with each other and their technologically advanced machines in a more proactive, rather than in a reactive manner, to ensure their own safety, and that of their passengers, the cargo and the intricate and expensive machines that they fly.

7 LIST OF REFERENCES

Aaker, D. A., Kumar, V. & Day, G. S. 1995. *Marketing research*. 5th edition. New York: Wiley.

Abbott, K. 1995. The evaluation of two CDU concepts and their effects on FMS training. *Eighth International Symposium on Aviation Psychology*, 24-27 April. Columbus, OH: Ohio State University.

Abbott, K. 2010. Operational use of flightpath management systems. *Flight Safety Foundation International Seminar*, 3-5 November, Milan.

Abraham, R. 1990. Managing safety in international operations. *Proceedings of the 43rd Annual International Air Safety Seminar,* 19-22 November, Rome.

Agresti, A. 1996. *An introduction to categorical data analysis*. New York: Wiley.

Airbus. 2009. *Flight crew training manual: Airbus A320 series*. [Online] Available from: http://www.smartcockpit.com/plane/airbus/A320/ [Accessed:10 March 2009].

Airbus. 2011a. *Flight crew operating manual, FCOM A330* V*olume 1. Systems description*. Toulouse: Airbus Training Centre.

Airbus. 2011b. *Flight crew training manual, FCTM A330/A340*. Toulouse: Airbus Training Centre.

Aliseda, A. 2006. *Abductive reasoning: logical investigations into discovery and explanation*. Dordrecht: Springer.

ALPA-SA. 2011. *Airline Pilots' Association of South Africa*. [Online] Available from: http://www.alpa.co.za/ [Accessed: 10 February 2011].

Altman, D. G. 1991. *Practical statistics for medical research.* New York: CRC.

American Psychological Association (APA). 1994. *Publication manual of the American Psychological Association*. 4th ed. Washington, DC: APA.

American Psychological Association (APA). 2002. *Ethical principles of psychologists and the code of conduct.* Washington, DC: APA.

Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26:32-46.

Andrews, D. H. & Thurman, R. A. 2000. An instructional systems development primer for aircrew training. In: O'Neil, H. F. & Andrews, D. H. (eds). *Aircrew training and assessment.* (pp. 110-152). New York: Erlbaum.

Andrich, D. 1978. Relationships between the Thurstone and Rasch approaches to item scaling. *Applied Psychological Measurement*, 2:449-460.

APA – see American Psychological Association

Argyris, C. 1957. *Personality and organization.* New York: Harper.

Aronson, J. 1991. Learning and education. In: Field, K., Cohler, B. & Wool, G. (eds). *Psychoanalytic perspectives. Emotions and behaviour* (pp. 65-97). Madison, WI: International Universities Press.

Arrindell, W. A. & Van der Ender, J. 1985. An empirical test of the utility of the observer-to-variables ratio in factor and components analysis. *Applied Psychological Measurement,* 9:165-175.

Ashcroft, K. & Foreman-Peck, L. 1994. *Managing teaching and learning in further and higher education.* Oxford: Routledge.

Ausink, J. A. & Marken, R. S. 2005. *Assessing the impact of future operations on trainer aircraft requirements*. Santa Monica, CA: Rand.

Australian Transport Safety Bureau. 2007. *How old is too old? The impact of ageing aircraft on aviation safety. ATSB transport safety report. Aviation research and analysis report. Final.* Sydney: Australian Government.

Babbie, E. R. 2010. *The practice of social research*. Belmont, CA: Wadsworth.

Bainbridge, L. 1983. Ironies of automation. *Automatica*, 19(6):775-779.

Ballinger, C. & Davey, C. 1998. Designing a questionnaire: an overview. *British Journal of Occupational Therapy*, 61:547-550.

Barker, D. 2011. Is aeronautical engineering getting too far ahead of itself? *World Air News*, April. [Online] Available from: http://www.airnews.co.za/ [Accessed: 21 May 2011].

Barnette, J. J. 2000. Effects of stem and Likert response option reversals on survey internal consistency: if you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, 60:361-372.

Bartlett, J. E., Kotrlik, J. W. & Higgins, C. C. 2001. Organizational research: determining sample size in survey research. *Information Technology, Learning, and Performance Journal,* 19(1):43-50.

Baum, A., Gatchel, R. J. & Schaeffer, M. A. 1983. Emotional, behavioural, and physiological effects of chronic stress at Three Mile Island. *Journal of Consulting and Clinical Psychology,* 51(4):565-572.

Bejar, I. I. 1983. Introduction to item response models and their assumptions. In: Hambleton, R. K. (ed.). *Applications of item response theory* (pp. 105-156). Vancouver: Educational Research Institute of British Columbia.

Bent, J. 1996. *Training for new technology*. Hong Kong: Cathay Pacific Airways.

Bentler, P. & Raykov, T. 2000. On measures of explained variance in nonrecursive structural equation models. *Journal of Applied Psychology*, 85(1):125-131.

Bergman, M. M. 2008. *Advances in mixed methods research.* New York: Sage.

Berliner, D. C. 2006. Educational psychology: searching for essence throughout a century of influence. In: Alexander, P. A. & Winne, P. H. (eds). *Handbook of educational psychology* (pp. 67-98). 2nd ed. New York: Routledge.

Biggs, J.B. 1987. *Student approaches to learning and studying.* Hawthorne: Australian Council for Educational Research.

Biggs, J.B. 1999. *Teaching for quality learning at university*. Sydney: Open University Press.

Biggs, J. B. & Moore, P. J. 1993. *The process of learning.* Sydney: Prentice-Hall.

Billiet, J. B. & McClendon, M. J. 2000. Modelling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modelling*, 7(4):608-628.

Billings, C. E. 1997. *Aviation automation: the search for a human-centred approach.* Mahwah, NJ: Erlbaum.

Boeing. 2009. *Statistical summary of commercial jet airplane accidents, worldwide operations 1959-2009.* Seattle, WA: Aviation Safety. Boeing Commercial Airplanes.

Boeing, 2010. *Boeing begins assembly of first intercontinental 747-8*. [Online] Available from: http://boeing.mediaroom.com/.html [Accessed: 10 October 2011].

Bonner, M. A. & Wilson, G. F. 2002. Heart rate measures of flight test and evaluation. *International Journal of Aviation Psychology*, (12):63–77.

Bott, J. P. & Svyantek, D. J. 2004. Organisational culture and organisational culture measures: an integrative review. In: Hersen, M. (ed.). *Comprehensive handbook of psychological assessment, industrial and organisational assessment. Volume 4.* (pp. 412-427). Aylesbury: Wiley.

Bradburn, N. 1983. Response effects. In: Rossi, P., Wright, J. & Anderson, A. (eds). *Handbook of survey research* (pp. 593-631). San Diego, CA: Academic Press.

Briere, D. & Traverse, P. 1993*. Airbus A320/A330/A340 electrical flight controls a family of fault-tolerant systems.* Toulouse: Aerospatiale.

Brown, F. G. 1983. *Principles of educational and psychological testing*. 3rd ed. New York: CBS.

Brown, B. L., Hendrix, S. B., Hedges, D. W. & Smith, T. B. 2012. *Multivariate analysis for the biobehavioral and social sciences. A graphic approach*. Hoboken, NJ: Wiley.

Burns, N. & Grove, S. K. 2005. *The practice of nursing research: conduct, critique and utilization.* London: Elsevier.

Butcher, H. J. 1966. *Statistical guides in educational research. No. 3.* Oxford: Manchester University Press.

Byrne, D. S. 1998. *Complexity theory and the social sciences: an introduction*. New York: Routledge.

CAA – see Civil Aviation Authority

Calaprice, A. 2005. *The new quotable Einstein*. Princeton, NJ: Princeton University Press.

Caro, P. W. 1988. Flight training and simulation. In: Wiener, E. L. & Nagel, D. C. (eds). *Human factors in aviation* (pp. 293-332). San Diego, CA: Academic Press.

Carston, R. 2002. *Thoughts and utterances. The pragmatics of explicit communication.* Malden: Blackwell.

Castellan, N. J. 1992. Shuffling arrays: appearances may be deceiving. *Behavior Research Methods,* 24(1):72-77.

Cattell, R. B. 1946. Personality structure and measurement. The determination and utility of trait modality. *British Journal of Psychology,* 36:159-174.

Cattell, R. B. 1966. The scree test for the number of factors. *Multivariate Behavioral Research*, 1:245-76.

Cattell, R. B. 1987. *Intelligence: its structure, growth, and action*. New York: Elsevier.

Cattell, R. B., Boyle, G. J. & Chant, D. 2002. The enriched behavioural prediction equation and its impact on structured learning and the dynamic calculus.

Humanities and Social Sciences Papers, Bond University. [Online] Available from: http://epublications.bond.edu.au [Downloaded: 10 November 2009].

Chambers, A. B. & Nagel, D. C. 1985. Pilots of the future: human or computer? *Computer*, 18(11), Nov:74-87.

Child, J. 1973. Predicting and understanding organization structure. *Administrative Science Quarterly*, 18(2), Jun:168-185.

Chung, W. V. 1996. *Ethnicity and organizational diversity: a study of social cognition and psychological climate perception*. Lanham, MD: University Press of America.

Civil Aviation Authority (CAA). 2011. *Acts and regulations*. [Online] Available from: http://www.caa.co.za/ [Accessed: 13 April 2011].

Clark, L. A. & Watson, D. 1995. Constructing validity: basic issues in objective scale development. *Psychological Assessment*, 7(3):309-319. [Online] Available from: http://0-www.sciencedirect.com.innopac.up.ac.za/ [Downloaded: 11 June 2009].

Clarke, A. C. 1984. *Profiles of the future: an inquiry into the limits of the possible*. New York: Holt, Rinehart, and Winston.

Clason, D. L. & Dormody, T. J. 2001. Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education*, 35(4):31-36.

Cobanoglu, C., Warde, B. & Moreo, P. J. 2001. A comparison of mail, fax and web-based survey methods. *International Journal of Market Research*, 43(4):441-454.

Cochran, W.G.1954. Some methods for strengthening the common χ^2 tests. *Biometrics*, 10(4):417-451.

Cockburn, B. 2007. Video and analysis. TAM A320 crash in Sao Paulo. *Flight Global, 18 July.* [Online] Available at: http://www.flightglobal.com/articles/ [Accessed: 21 March 2009].

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences.* 2nd ed. Newark, NJ: Erlbaum.

Cohen, J. 1992. Statistical power analysis. *Current Directions in Psychological Science*, 1(3):98-101.

Cohen, J., Cohen, P., West, S. G. & Alken, L. S. 2003. *Applied multiple regression/correlation analysis for the behavioral sciences.* 3rd ed. Newark, NJ: Erlbaum,

Cohen, B. H. & Lea, R. B. 2004. *Essentials of statistics for the social and behavioural sciences.* Newark NJ: Wiley.

Comrey, A. L. & Lee, H. B. 1992. *A first course in factor analysis.* 2nd ed. Hillside, NJ: Erlbaum.

Cooper, D. R. & Schindler, P. S. 2003. *Business research methods.* 8th edition. New York: McGraw-Hill/Irwin.

Corston, R. & Colman, A. 2003. *A crash course in SPSS for windows*. 2nd ed. Oxford: Blackwell.

Cortina, J. M. 1993. What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1):98-104.

Costa, P. T. & McCrae, R. R. 1992. *Revised neo personality inventory (neo-pi-r) and neo five-factor inventory (neo-ffi) manual*. Miami, FL: Psychological Assessment Resources.

Creswell, J. W. 2002. *Research design. Quantitative, qualitative and mixed methods approaches.* 2nd ed. Nebraska-Lincoln, NE: Sage.

Cronbach, L. J. 1942. Studies of acquiescence as a factor in the true-false test. *Journal of Educational Psychology*, 33:401-415.

Cronbach, L. J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297-334.

Davis, W. D., Fedor, D. B., Parsons, C. K. & Herold, D. M. 2000. The development of self-efficacy during aviation training. *Journal of Organizational Behavior,* 21(8):857-871.

Dahlstrom, N. & Nahlinder, S. 2006. A comparison of two recorders for obtaining in-flight heart rate data. *Applied Psychophysiology and Biofeedback*, (31):273–279.

Degani, A., Shafto, M. & Kirlik, A.1995. On the types of modes in human-machine interfaces. In: Jensen, R. S. & Rakovan, L. A. (eds). *Proceedings of the Eighth International Symposium on Aviation Psychology,* Ohio, 24-27 April:959-965.

Dekker, S. W. A. & Johansson, B. 2000. *JAR-FCL and pilot knowledge needs*. Linköping, Sweden: Swedish Centre for Human Factors in Aviation, Linköping University.

Denison, D. R. 1996. What is the difference between organizational culture and organizational climate? A native’s point of view on a decade of paradigm wars. *Academy of Management Review*, 21(3):619-654.

Denzin, N. K. & Lincoln, Y. S. 2005a. Introduction: the discipline and practice of qualitative research. In: Denzin, N. K. & Lincoln, Y. S. (eds.) *Strategies of qualitative inquiry* (pp. 125-148). Thousand Oaks, CA: Sage.

Denzin, N. K. & Lincoln, Y. S. 2005b. *The Sage handbook of qualitative research.* New York: Sage.

Desler, G. 2002. *Human resource management*. 9th ed. New York: Prentice-Hall.

De Vaus, D. A. 2003. *Research design in social research*. Thousand Oaks, CA: Sage.

DeVellis, R. F. 2003. *Scale development: theory and applications*. Newbury Park, CA: Sage.

Dole, C. E. 1989. *Flight theory for pilots*. Englewood, Chicago: Jeppesen Sanderson.

Dobson, A. J. & Barnett, A. G. 2008. *Introduction to generalized linear models.* 3rd ed. Boca Raton, FL: Chapman and Hall/CRC.

Drucker, P. F. 1946. *The concept of the corporation.* New York: Day.

Dutch Safety Board. 2009. *Report of Turkish B737-800 accident, Schipol*. [Online] Available from: http://www.onderzoeksraad.nl/ [Accessed:10 March 2009].

Ellis, P. D. 2010. *The essential guide to effect sizes: an introduction to statistical power, meta-analysis and the interpretation of research results*. London: Cambridge University Press.

Embretson, S. & Reise, S. 2000. *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Entwistle, N. & McCune, V. 2004. The conceptual basis of study strategy inventories. *Educational Psychology Review*, 16(4), 325-345. [Online] Available from: https://commerce.metapress.com/ [Downloaded: 25 January 2009].

FAA – see Federal Aviation Administration

Federal Aviation Administration (FAA). 1996. *FAA Human factors team report: the interface between flight crews and modern flight deck systems.* Washington, DC: FAA.

Feeney, A. & Heit, E. 2007. *Inductive reasoning: experimental, developmental, and computational approaches*. New York: Cambridge University Press.

Feyerabend, P. K. 1985. *Realism, rationalism and scientific method.* New York: Press Syndicate.

Field, A. 2005. *Discovering statistics using SPSS (and sex, drugs and rock 'n roll).* 2nd ed. London: Sage.

Field, A. 2009. *Discovering statistics using SPSS.* New York: Sage.

Fishbein, M. & Ajzen, I. 2001. *Belief, attitude, intention and behaviour: an introduction to theory and research.* Reading, MA: Addison-Wesley.

Fiske, D. W. 2009. Problems in measuring personality. In: Wepman, J. & Heine, R. W. (eds). *Concepts of personality* (pp. 56-78). Chicago: Aldine.

Fleiss, J. L., Levin, B. A. & Paik, M. C. 2003. *Statistical methods for rates and proportions.* New York: Wiley.

Funk, K. & Lyall, B. 2000. *A comparative analysis of flight decks with varying levels of automation. Final report*. Washington, DC: FAA. [Online] Available at: http://www.researchintegrations.com/ [Accessed:10 February 2009].

Gegax, D., Gerking, S. & Schulze, W. 1991. Perceived risk and marginal value of safety. *Review of Economics and Statistics*, 2(1):589-597.

Gentle, J. E. 2007. *Random number generation and Monte Carlo methods*. New York: Springer.

Gerbing, D. W. & Anderson, J. C. 1988. An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25(2):186-192.

Gliem, J. & Gliem, R. 2003. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. *Midwest Research to Practice Conference in Adult, Continuing and Community Education,* 26-28 September*.* Columbus, OH.

Glorfeld, L. W. 1995. An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, 55(3):377-393.

Go, T. H., Bürki-Cohen, J. & Soja, N. N. 2003. The effects of enhanced hexapod motion on airline pilot recurrent training and evaluation. *American Institute of Aeronautics and Astronautics Conference Paper*, 2000-4296. 11-14 August. Austin, TX.

Goldstein, H. & Wood, R. 1989. Five decades of item response modelling. *British Journal of Mathematical and Statistical Modelling*, 42:139-167.

Gorsuch, R. L. 1983. *Factor analysis*. Hillsdale, NJ: Erlbaum.

Gorsuch, R. L. 1997. Exploratory factor analysis: its role in item analysis. *Journal of Personality Assessment,* 68(3):532-560.

Govindarajulu, Z. 2001. *Statistical techniques in bioassay*. 2nd ed. Basel: Reinhardt.

Gravetter, F. J. & Wallnau, L. B. 2008. *Statistics for the behavioural sciences.* New York: Wadsworth.

Green, S. B. & Salkind, N. J. 2008. *Using SPSS for Window and Macintosh: analyzing and understanding data.* 5th ed. Upper Saddle River, NJ: Pearson Prentice Hall.

Hakala, C. 2009. *AP psychology. Proven tools to help you score higher.* Berkeley, CA: Kaplan.

Hambleton, R. K. & Rogers, H. J. 1989. Detecting potentially biased test items: comparison of IRT area and Mantel-Haenszel methods. *Applied Measurement in Education,* 2(4):313-334.

Hardesty, D. M. & Bearden, W. O. 2004. The use of expert judges in scale development, implications for improving face validity of measures of unobservable constructs. *Journal of Business Research,* 57:98-107.

Harman, H. H. & Yoichiro, F. 1966. Resolution of the Heywood case in the minres solution. *Psychometrika*, 31(4):563-571.

Harrison, J. P. 2000. *Mastering the sky: a history of aviation from ancient times to the present.* New York: Sarpedon.

Haworth, T. J. 1996. *Psychological research: innovative methods and strategies.* London: Routledge.

Hayton, J. C., Allen, D. G. & Scarpello, V. 2004. Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis*. Organizational Research Methods, 7*(2):191-205.

Hellreigel, D. & Slocum Jr., J. W. 1974. Organizational climate: measures, research and contingencies. *Academy of Management Journal*, 17:255-281.

Helmreich, R. L. 1987. Flight crew behaviour. *Social behaviour*, 2:63-72.

Helmreich, R. L. 2002. Managing threat and error in aviation and medicine. In: *Proceedings of the Third LOSA Week* (ICAO):15-22. Dubai, United Arab Emirates, October 14-16.

Hendrickson, A. E. & White, P. O. 1964. Promax: a quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology*, 17(1):65-70.

Ho, R. 2006. *Handbook of univariate and multivariate data analysis and interpretation with SPSS*. New York: Chapman and Hall.

Holmes, F. 2005. Comparison of the performance of nonparametric and parametric MANOVA statistics when assumptions are violated. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 1(1):27-38.

Holton, E. F., Bates, R. A., Seyler, D. & Carvalho, M. 2000. Toward a construct validation of a transfer climate instrument. *Human Resource Development Quarterly*, 8(1):95-113.

Horn, J. L. 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185.

Hosmer, D. W. & Lemeshow, S. 2000. *Applied logistic regression*. 2nd ed. New York: Wiley.

Howell, W. C. & Fleishman, E. A. 1982. *Information processing and decision making.* New York: Erlbaum.

Hradecky, S. 2011. *Report: Air Canada A333 near Montreal on Apr 30th 2008, inflight engine fuel starvation*. [Online] Available from: http://www.avherald.com [Accessed:9 March 2011].

IATA – see International Air Transport Association

International Civil Aviation Organisation. 2001. *Aircraft accident and incident investigation. International standards and recommended practices*. *Annex 13*.

Montreal, Canada. [Online] Available from: http://www.iprr.org/manuals/Annex13.html [Accessed:18 May 2011].

International Air Transport Association, 2012. *Standard schedules information manual (ssim)*. Montreal, Canada. [Online] Available from: http://www.iata.org/ps/publications/Pages/ssim.aspx [Accessed:28 March 2012].

International Civil Aviation Organisation. 2011. *Convention on international civil aviation*. Montreal, Canada. [Online] Available from: http://www.icao.int/icaonet/dcs/7300.html [Accessed:18 March 2011].

Ishida, T. & Kanda, N. 1999. Analysis of aircraft accidents by means of variation tree. *International Journal of Aviation Psychology*, 22(2):109-120. [Online] Available from: http://dx.doi.org/ [Downloaded: 22 March 2009].

James, L. R., Jones, A. P. & Ashe, D. K. 1990. The meaning of organizations: the role of cognition and values. In: Schneider, B. (ed.). *Organizational climate and culture* (pp. 382-424). San Francisco, CA: Jossey-Bass.

James, M., McClumpha, A., Green, R., Wilson, P. & Belyavin, A. 1991. Pilot attitudes to flight deck automation. *Proceedings of the Royal Aeronautical Society Conference on Human Factors on Advanced Flight Decks,* London, 9-10 March:130-158.

Johnston, N., Fuller, R. & McDonald, N. 1995, *Aviation psychology: training and selection. Proceedings of the 21st Conference of the European Association for Aviation Psychology (EAAP),* Avebury Aviation, Aldershot.

Josephson, J. R. & Josephson, S. G. 1996. *Abductive inference: computation, philosophy, technology*. New York: Cambridge University Press.

Kaiser, H. F. 1961. A note on Guttman's lower bound for the number of common factors. *British Journal of Statistical Psychology*, 14(1):1-20.

Kaminski-Morrow, D. 2009. Nine killed as Turkish 737 breaks up during Amsterdam landing. *Flight Global,* 25 February. [Online] Available from: http://www.flightglobal.com/ [Accessed: 22 March 2009].

Kalton, G. 1999. *Introduction to survey sampling*. Thousand Oaks, CA: Sage.

Karlsson, C. 2008. *Researching operations management.* New York: Taylor and Francis.

Katz, D. & Khan, R. L. 1966. *The social psychology of organizations.* New York & London: Wiley.

Keane, P. J. 2005. *Emerson, romanticism, and intuitive reason: the transatlantic light of all our day*. Columbia, MO: University of Missouri Press.

Keeter, S., Miller, C., Kohut, A., Groves, R. M. & Presser, M. 2000. Consequences of reducing non-response in a national telephone survey. *Public Opinion Quarterly*. [Online] Available from: http://pollcats.net/downloads/nonresponse.final.pdf [Downloaded: 18 September 2009].

Kendall, M. G. & Stuart, A. 1963. *The advanced theory of statistics*, *Vol. 1.* 2nd ed. London: Charles Griffith.

Kolb, D. A. 1976. *Learning style inventory: technical manual.* Boston, MA: McBer.

Koonce, J. M. 2003. *Human factors in the training of pilots*. London: Taylor & Francis.

Kozlowski, S.W. & Hults, B. M. 1987. An exploration of climates for technical updating and performance. *Personnel Psychology*, 40(1):539-563.

Kristovics, A. (A.kristovics@uws.edu.au) 2010. Balanced item pool. [Email to:] Vermeulen, L. P. (lvermeul@tiscali.co.za) 30 March 2010.

Krus, D. J. 2006. Reliability and homogeneity. *Journal of Visual Statistics*, 6:35-47.

Landis, J. R. & Koch, G. G. 1977. The measurement of observer agreement for categorical data, *Biometrics,* 33(1):159-174. [Online] Available from: http://www.jstor.org.innopac.up.ac.za/ [Downloaded: 10 July 2009].

Lawley, D. N. 1939. A generalization of Fisher's IX test. *Biometrika,* 30:180-187.

Lawshe, C. H. 1975. A quantitative approach to content validity. *Personnel Psychology,* 28:563-575.

Ledyard, R. T. 1966. Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31(3):279-311.

Leech, N. L. 2004. Enhancing the interpretation of “significant” findings: the role of mixed-methods research. *Qualitative Report,* (9)4:760-810.

Leech, N. L., Barrett, K. C. & Morgan, G. A. 2005. *SPSS for intermediate statistics. Use and interpretation.* 2nd ed. Mahwah, NJ: Erlbaum.

Leibold, M., Probst, G. & Gibbert, M. 2005. *Strategic management in the knowledge economy*. *New approaches and business applications*. 2nd ed. Erlangen, Germany: Publicis and Wiley.

Likert, R. 1932. A technique for the measurement of attitudes. *Archives of Psychology*, 22(140):34-53.

Likert, R. 1958. Measuring organizational performance. *Harvard Business Review,* March-April:41-51.

Lilliefors, H. W. 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318):339-402.

Loevinger, J. 1957. Objective tests as instruments of psychological theory. *Psychological Reports,* 3:635-694.

Long, J. S. 1997. *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage.

Lowy, J. 2009. Pilot training, fatigue issues in Buffalo crash. *Washington Post*, 11 May. [Online] Available from: http://www.washingtonpost.com/ [Accessed:12 May 2009].

Lyall, B. & Funk, K. 1998. Flight deck automation issues. In: Scerbo, M.W. & Mouloua, M. (eds). *Automation technology and human performance. Current research and trends* (pp.288-296). Norfolk, VA: Routledge.

MacCallum, R. C., Widaman, K. F., Zhang, S. & Hong, S. 1999. Sample size in factor analysis. *Psychological Methods*, 4(1):82-99.

Machin, M. A. & Fogarty, G. J. 2003. Perceptions of training-related factors and personal variables as predictors of transfer implementation intentions. *Journal of Business and Psychology*, 18(1):51-71.

Magnusson, S. 2002. Similarities and differences in psychophysiological reactions between simulated and real air-to-ground missions. *International Journal of Aviation Psychology*, (12):49–61.

Mardia, K. V., Kent, J. T. & Bibby J. M. 1979. *Multivariate analysis*. New York: Academic Press.

Masters, J. R. 1974. The relationship between number of response categories and reliability of Likert type questionnaires. *Journal of Educational Measurement,* 11(1):49-53. [Online] Available from: http://0-www.jstor.org.innopac.up.ac.za/ [Downloaded: 18 February 2010].

McManus, I. C., Keeling, A. & Paice, E. 2004. Stress, burnout and doctors' attitudes to work are determined by personality and learning style: a twelve year longitudinal study of UK medical graduates. *BMC Medicine*, 2(29). [Online] Available from: http://www.biomedcentral.com/ [Accessed: 29 February 2009].

Meijer, R. R. & Baneke, J. J. 2004. Analysing psychopathology items: a case for nonparametric response theory modelling. *Psychological Methods*, 9(3):354-368.

Meister, D. 1999. Measurement in aviation systems. In: Garland, J. D., Wise, J. A. & Hopkin, V. D. (eds). *Handbook of aviation human factors* (pp. 33-64). New York: Erlbaum.

Milmo, D. & Webb, T. 2010. Rolls-Royce faces backlash in US over Qantas A380 engine failure. *Guardian*, 5 November:10.

Mitchell, J., Vermeulen, L. P. & Naidoo, P. 2009. Flying glass: a qualitative analysis of pilot perceptions of automated flight-decks after 20 years. *International Journal of Applied Aviation Studies*, 9(1):13-29. Oklahoma City, OK: FAA Academy.

Moon, J. A. 2004. *A handbook of reflective and experiential learning*. *Theory and practice*. New York: Routledge/Falmer.

Moore, G. 2003. Moore optimistic on Moore's Law. *Intel Architecture & Silicon Technology.* [Online] Available from: http://www.intel.com/technology/ [Accessed: 26 May 2011].

Moore, P. J., Lehrer, H. R. & Telfer, R. A. 1997. Instructor perspectives on learning in aviation. In: Jensen, R. & Rakovan, L. (eds). *Proceedings of the Ninth Annual International Symposium on Aviation Psychology*. Columbus, OH: Ohio State University. [Online] Available from: http://www.emeraldinsight.com/Insight/ [Downloaded: 30 March 2009].

Moore, P. J., Po, T., Lehrer, H. R. & Telfer, R. A. 2001. Quality training and learning in aviation: problems of alignment. *Journal of Air Transportation Worldwide,* 6(1). [Online] Available from: http://ntl.bts.gov/lib/000/700/744/atww_6-1_Moore-1.pdf [Downloaded: 10 April 2009].

Morgan, A. G. & Griego, O. V. 1998. *Easy use and interpretation of SPSS for Windows: answering research questions with statistics*. London: Erlbaum.

Morgan, G. A., Leech, N. L., Gloeckner, G. W. & Barrett, K. C. 2007. *SPSS for introductory statistics: use and interpretation.* 3rd edition. London: Erlbaum.

Moses, L. N. & Savage, I. 1989. The effect of airline pilot characteristics on perceptions of job safety risks. *Journal of Risk and Uncertainty*, 2(4):335-351.

Mosier, K. L., Sethi, N., McCauley, S., Khoo, L. & Orasanu, J. M. 2007. What you don't know can hurt you: factors impacting diagnosis in the automated cockpit. *Human Factors: Journal of the Human Factors and Ergonomics Society,* 49(2):300-310.

Mouloua, M., Gilson, R. D. & Koonce, J. 1997. Automation, flight management and pilot training: issues and considerations. In Telfer, R. A. & Moore, P. J. (eds). *Aviation training: learners, instruction and organisation* (pp.110-135). Aldershot: Avebury Aviation.

Mouton, J. & Marais, H. C. 1996. *Basic concepts in the methodology of the social sciences.* 5th ed. Pretoria: HSRC.

Muijs, D. 2004. *Doing quantitative research in education with SPSS*. New York: Sage.

Myers, R. H., Montgomery, D. C. & Cook, C. M. A. 2009. *Response surface methodology.* New York: Wiley.

Naidoo, P. 2008. *Airline pilots' perceptions of advanced flight deck automation.* Unpublished master's thesis, University of Pretoria, Pretoria, South Africa.

National Transportation Safety Board (NTSB). 2009. *Accident and incident report for Part 121 Operators.* [Online] Available from: http://www.ntsb.gov/ntsb/AVIATION/ [Downloaded: 17 March 2009].

Neal, A., Griffin, M. A. & Hart, P. M. 2000. The impact of organizational climate on safety climate and individual behaviour. *Safety Science*, 34:99-109.

Nelson, R. R. 1959. The simple economics of basic scientific research. *Journal of Political Economy*, 67(3):297-306.

Netemeyer, R. G., Bearden, W. O. & Sharma, S. 2003. *Scaling procedures, issues and applications*. Thousand Oaks, CA: Sage.

Nunnally, J. C. & Bernstein, I. H. 1994. *Psychometric theory.* 3rd edition. New York: McGraw-Hill.

Oaksford, M., Chater, N. & Hahn, U. 2008. Human reasoning and augmentation: the probabilistic approach. In: Adler, J. E. & Rips, J. (eds). *Reasoning. Studies of human inference and its foundations* (pp. 383-405). New York: Cambridge University Press.

O'Connor, B. P. 2000. SPSS and SAS programs for determining the number of components: using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments & Computers,* 32, 396-402. [Online] Available from: https://people.ok.ubc.ca/brioconn/nfactors/rawpar.sps [Accessed: 20 July 2011].

O'Hare, D., Wiggins, M., Batt, R. & Morrison, D. 1994. Cognitive failure analysis for aircraft accident investigation. *Ergonomics,* 37(11):185-197.

Olney, S. L. 1996. *Unions in a changing world: problems and prospects in selected industrialized countries.* London: International Labour Office.

Oosterhof, A. C. 1976. Similarity of various item discrimination indices. *Journal of Educational Measurement*, 13(2):145-155.

Osborne, J. W. & Costello, A. B. 2004. Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research & Evaluation*, 9(11). [Online] Available from: http://PAREonline.net/ [Accessed: 16 July 2011].

Panda, S. 2003. *Planning and management in distance education.* New York: Routledge.

Parasuraman, R. & Byrne, E. A. 2002. Automation and human performance in aviation. In: Tsang, P. S. & Vidulich, M. A. (eds.) *Principles and practices of aviation psychology* (pp.311-356). Washington, DC: Erlbaum.

Parasuraman, R. & Riley, V. 1997. Humans and automation: use, misuse, disuse, abuse. *International Journal of Aviation Psychology,* 39(2):230-253.

Pasztor, A. 2009. Captain's training faulted in air crash that killed 50. *Wall Street Journal,* 11 May.

Patrick, J. 2002. Trends and contexts of pilot training. In: Tsang, P.S. & Vidulich, M.A. (eds). *Principles and practices of aviation psychology* (pp.115-134). Washington, DC: Erlbaum.

Pepper, S. C. 1970. *World hypotheses. Prolegomena to systematic philosophy and a complete survey of metaphysics.* Los Angeles: University of California Press.

Pett, M. A., Lackey, N. R. & Sullivan, J. J. 2003. *Making sense of factor analysis.* Thousand Oaks, CA: Sage.

Pitfield, D. E., Caves, R. E. & Quddus, M. A. 2010. Air line strategies for aircraft size and air line frequency with changing demand and competition: a simultaneous equations approach for traffic on the North Atlantic. *Journal of Air Transport Management*, 16 (3):151-158.

Pietarinen, A. 2006. On historical pragmatics and Peircean pragmatism. *Linguistics and the Human Sciences*, 2(1):123-143.

Pohlman, L. D. & Fletcher, J. D. 1999. Aviation personnel selection and training. In: Garland, J. D., Wise, J. A. & Hopkin, V. D. (eds). *Handbook of aviation human factors* (pp. 277-308). New York: Erlbaum.

Pololi, L. & Price, J. 2000. Validation and use of an instrument to measure the learning environment as perceived by medical students. *Teaching and Learning in Medicine*, 12(4):201-237. [Online] Available from: http://www.informaworld.com/smpp/ [Accessed: 10 May 2009].

Poprawa, S. 2011. Flight control automation. *SA Flyer*, December. [Online] Available from: http://www.saflyer.co.za/ [Accessed: 20 January 2012].

Ramsden, P. 1992. *Learning to teach in higher education*. New York: Routledge.

Rasmussen, J. 1990. Human error and the problem of causality in analysis of accidents. *Philosophical Transactions of the Royal Society B*, 327:449-462.

Rencher, A. C. 2002. *Methods of multivariate analysis*. 2nd ed. New York: Wiley.

Research Integrations Inc. 2007. *Flight deck automation issues.* [Online] Available from: http://www.flightdeckautomation.com/fdai.aspx [Accessed: 11 April 2009].

Reynolds, J. 2001. *Succeeding in college: study skills and strategies.* New York: Prentice Hall.

Ribbens, W. B. & Mansour, N. P. 2003. *Understanding automotive electronics.* London: Elsevier.

Rigner, J. & Dekker, S. 2000. Sharing the burden of flight deck automation training. *International Journal of Aviation Psychology*, 10(4):317-326. [Online] Available from: EBSCOHost EJS Content Distribution: http://dx.doi.org/10.1207/S15327108IJAP1004_1 [Accessed:19 February 2009].

Risukhin, V. 2001. *Controlling pilot error. Automation.* New York: McGraw-Hill.

Robbins, S. P., Odendaal, A. & Roodt, G. 2004. *Organisational behaviour. Global and Southern African perspectives.* Cape Town: Maskew Miller Longman.

Roff, S. 2005. The Dundee Ready Educational Environment Measure (DREEM), a generic instrument for measuring students' perceptions of undergraduate health professions curricula. *Medical Teacher*, 27(4):322-325. [Online] Available from: Informaworld: http://0-www.informaworld.com.innopac.up.ac.za/smpp/ [Downloaded: 15 April 2009].

Rogers, C. R. & Skinner, B. F. 1956. Some issues concerning the control of human behaviour. *Science,* 124:1057-1066.

Rogers, J., Monteiro, F. M. & Nora, A. 2008. Toward measuring the domains of mentoring. *Fam Med*, 40(4):259-263.

Rolfe, J. M. & Staples, K. J. 1986. *Flight simulation.* Cambridge: Press Syndicate.

Rosenow, E. C. 2003. *The art of living: the art of medicine the wit and wisdom of life and medicine: a physician's perspective*. Victoria, Canada: Trafford.

Rosenthal, R. 1994. Science and ethics in conducting, analyzing, and reporting psychological research. *Psychological Science,* 5(3):127-136.

Rouiller, J. A. & Goldstein, I. L. 1993. The relationship between organizational transfer climate and positive transfer of training. *Human Resource Development Quarterly*, 4(1):377-390.

Rouse, W. B. & Morris, N. M. 1987. Conceptual design of human error tolerant interface for complex engineering systems. *Automatica,* 23(2):231-235.

SACAA – see South African Civil Aviation Authority

Sarter, N. B. 1996. Human errors are symptoms of a mismatch between pilots, machines, and the operating environment. *ICAO Journal,* 51(8):9-15.

Sarter, N. B. & Woods, D. D. 1994. Pilot interaction with cockpit automation: an experimental study of pilots' model and awareness of the flight management system. *International Journal of Aviation Psychology,* 4(1):1-25. [Online] Available from: http:// ebscohost.com.innopac.up.ac.za [Downloaded: 10 February 2009].

Saunders, M., Lewis, P. & Thornhill, A. 2007. *Research methods for business students.* 4th ed. Harlow: Pearson.

Sawilowsky, S. S. 2003. "You think you've got trivials?". *Journal of Modern Applied Statistical Methods,* 2(1):218-225.

Schaap, P. 2000. *Die ontwikkeling van 'n psigometriese instrument om die leerbenaderings van volwassenes te bepaal*. Unpublished DCom thesis, University of Pretoria, Pretoria, South Africa. [Online] Available from: http://upetd.up.ac.za/thesis/available/etd- 01192007-162401 [Downloaded: 15 June 2009].

Schaap, P. (pieter.schaap@up.ac.za) 2010. RE: Research proposal. [Discussion group:] freudian@telkomsa.net. 17 March 2010.

Schaap, P. 2011. The differential item functioning and structural equivalence of a nonverbal cognitive ability test for five language groups. *SA Journal of Industrial Psychology/SA Tydskrif vir Bedryfsielkunde*, 37(1) Art. #881, 16 pages. doi:10.4102/ sajip.v37i1.881.

Schein, E. H. 2004. *Organizational culture and leadership*. 3rd ed. San Francisco: Jossey-Bass.

Schreiner, E. 2010. *Specific principles of good scientific research*. [Online] Available from: http://www.ehow.co.uk/ [Accessed: 19 March 2011].

Schriesheim, C. A., Powers, K. J., Scandura, T. A., Gardiner C. C. & Lankau, M. J. 1993. Improving construct measurement in management research: comments and a quantitative approach for assessing the theoretical content adequacy of paper-and-pencil survey-type instruments. *Journal of Management*, (19)2:385-417.

Schuman, H. & Presser, S. 1996. *Questions and answers in attitude surveys: experiments on question form, wording and context*. London: Sage.

Schutte, P.C. 1998. Why humans are necessary on the flight deck: supporting the requisite role of the pilot. In: Scerbo, M.W. & Mouloua, M. (eds.) *Automation technology and human performance. Current research and trends* (pp. 318-320). Norfolk, VA: Routledge.

Seabrook, M. A. 2004. Clinical students' initial reports of the educational climate in a single medical school. *Medical Education*, 38(1):659-669.

Sherman, P. J. 1997. *Aircrews' evaluations of flight deck automation training and use: measuring and ameliorating threats to safety. Technical report.* Austin, TX: University of Texas.

Sherman, P. J., Helmreich, R. L. & Hines, W. E. 1995. The risks of automation: some lessons from aviation, and implications for training and design. *Proceedings of the workshop on Human error and Systems Development*, Glasgow, Scotland, 5-7 March: 1-7.

Simon, C., Little, P., Birtwistle, J. & Kendrick, T. 2003. A questionnaire to measure satisfaction with community services for informal carers of stroke patients: construction and initial piloting. *Health & Social Care in the Community*, 11(2):129-137.

Singh, I. L., Sharma, H. O. & Singh, A. L. 2005. Effect of training on workload in flight simulation task performance. *Journal of the Indian Academy of Applied Psychology,* 31(1-2):83-91.

Smith, G. M. & Dismukes, R. K. 2000. *Facilitation and debriefing in aviation training and operations*. London: Ashgate.

South African Airways (SAA). 2007. *Flight Operations Manual (FOM).* Johannesburg: Jeppesen Sanderson.

South African Civil Aviation Authority (SACAA). 2011a. *Publications*. *Licencing and examinations*. [Online] Available from: http://www.caa.co.za/ [Accessed: 20 April 2011].

Spector, M. J. & Muraida, D. J. 1997. Automating instructional design. In: Dijkstra, S., Robert, D. T., Schott, F., Seel, M. & Tennyson, D. (eds). *Instructional design: solving instructional design problems* (pp. 59-82). New York: Erlbaum.

Stoker, D. J. 1981. *Steekproefneming in die praktyk*. Pretoria: Van Schaik.

Strachan, I. W. 2011. *Collimation – diagram and real sim*. Wikimedia. [Online] Available from: http://commons.wikimedia.org. [Accessed: 18 March 2011].

Streiner, D. L. 2003. Being inconsistent about consistency: when coefficient alpha does and doesn’t matter. *Journal of Personality Assessment*, 80(1):99-103. [Online] Available from: http://0-web.ebscohost.com.innopac.up.ac.za/ehost/pdf [Downloaded: 13 June 2009].

Stevens, B. L. & Lewis, F. L. 2003. *Aircraft control and simulation*. New York: Wiley.

Stevens, S. S. 1946. On the theory of scales and measurement. *Science,* 103(2684):677-680.

Sullivan, L. E. 2009. *The Sage glossary of the social and behavioral sciences*. Thousand Oaks, CA: Sage.

Tabachnick, B. G. & Fidell, L. S. 2007. *Using multivariate statistics.* 5th edition. Boston, MA: Allyn & Bacon.

Tagiuri, R. & Litwin, G. 1968. *Organizational climate: explorations of a concept*. Boston, MA: Harvard Business School.

Taylor, H. L. & Emanuel, T. W. 2000. A civil aviation view of aircrew training. In: O'Neil, H .F. & Andrews, D. H. (eds). *Aircrew training and assessment* (pp. 17-36). New York: Erlbaum.

Teddlie, C. & Tashakkori, A. 2008. *Foundations of mixed methods research. Integrating quantitative and qualitative approaches in the social and behavioral sciences.* New York: Sage.

Telfer, R. A. & Moore, P. J. 1997. Introduction. The roles of learning, instruction and the organisation in aviation training. In: Telfer, R. A. & Moore, P. J. (eds). *Aviation training: learners, instruction and organisation* (pp.103-115). Aldershot: Avebury Aviation.

Telfer, R. A., Moore, P. J. & Farquharson, D. 1996. Pilot learning and conversion training. In: Hayward, B. J. & Lowe, A. R. (eds). *Applied aviation psychology. Achievement, change and challenge* (pp. 85-117). Brookfield, WI: Avebury Aviation.

Thorndike, E. L. 2007. *Educational psychology*. New York: Read Country Books.

Thurstone, L. L. 1947. *Multiple-factor analysis: a development and expansion of the vectors of mind.* Chicago: University of Chicago Press.

Tooley, M. 2007. *Aircraft digital electronic and computer systems: principal operation and maintenance.* Oxford: Elsevier.

Tracey, J. B. & Tews, M. J. 2005. Construct validity of a general training climate scale. *Organizational Research Methods,* 8(4):353-375.

Uebersax, J. S. 2006. Likert scales: dispelling the confusion. *Statistical methods for rater agreement.* [Online] Available from: http://john-uebersax.com/stat/likert.htm [Accessed: 13 March 2010].

Velicer, W. F. 1976. Determining the number of components from the matrix of partial correlations. *Psychometrika*, 31:321-327.

Vermeulen, L. P. (lvermeul@tiscali.co.za) 2011. Discussion of sampling method. [Email to:] Naidoo, P. (Freudian@telkomsa.net) 21 March 2011.

Vermeulen, L. P. 2009. Flight instructors' perceptions of pilot behaviour related to gender. *SA Journal of Industrial Psychology*, 35(1):128-136.

Voller, V. R. & Agel-Porte, F. 2002. Moore's law and numerical modelling. *Journal of Computational Physics*, 179:698-703.

Wald, M. L. 2009. Clues point to speed issues in Air France crash. *New York Times,* 4 June. [Online] Available from: http://www.nytimes.com/2009/06/05/world/europe/05plane.html [Accessed: 30 January 2010].

Walker, E. M. 1996. Questionnaire design in practice. *British Journal of Therapy and Rehabilitation,* 3(4):229-233.

Walker, D. A. 2003. Converting Kendall's Tau for correlational or meta-analytic analyses. *Journal of Modern Applied Statistical Methods*, 2(2):525-530.

Walsh, D. J. 1994. *On different planes: an organizational analysis of cooperation and conflict among airline unions*. Ithaca, NY: Cornell University.

Walters, S. J. C. 2002. *Study guide for Aviation Management 320. Aviation Safety Management.* Pretoria: Department of Human Resources Management, University of Pretoria.

Warner, R. M. 2008. *Applied statistics: from bivariate through multivariate techniques*. New York: Sage.

Warren, H. 2004. *Learning and teaching theory for engineering academics.* Loughborough: Engineering Subject Centre. [Online] Available from: http://www.engsc.ac.uk/ [Accessed: 1 May 2009].

Watkins, M. W. 2006. Determining parallel analysis criteria. *Journal of Modern Applied Statistical Methods,* 5 (2):344-346.

Welman, J. C. & Kruger, S. J. 1999. *Research methodology for the business and administrative sciences.* Halfway House, South Africa: Thomson International.

Wickens, C. D. 2000. *Engineering psychology and human performance.* 3rd ed. New York: Harper Collins.

Wiener, E. L. 1993. *Crew coordination and training in the advanced cockpit. Cockpit Resource Management.* San Diego, CA: Academic Press.

Wiener, E. L. 1998. Cockpit automation. In: Wiener, E. L. & Nagel, D. C. (eds). *Human factors in aviation* (pp. 85-117). San Diego, CA: Academic Press.

Wilson, A. M. & Weston, G. 1989. Application of airline pilots' hours to junior doctors. *Bio Medical Journal*, 299:779-781.

Wilson, N. & McClean, S. 1994. *Questionnaire design: a practical introduction.* Belfast: University of Ulster.

Wise, J. A., Tilden, D. S., Abbott, D., Dyck, J. & Guide, P. 1994. Managing automation in the cockpit. *Paper presented at the Flight Safety Foundation 47th Annual International Air Safety Seminar*, Oct 31 – Nov 3, Lisbon: Portugal.

Yeh, Y. C. 1996. *Triple-triple redundant 777 primary flight computer*. Seattle, WA: Boeing Commercial Airplane Group.

Zametti, F. 2008. *Practical DWR 2 projects.* New York: Springer. [Online] Available from: http://www.springerlink.com [Downloaded: 2 June 2009].

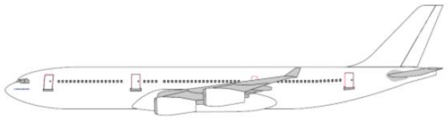
Zeller, R. A. & Carmines, E. G. 1980. *Measurement in the social sciences: the link between theory and data*. New York: Cambridge University Press.

Zwick, R. 1985. Nonparametric one-way multivariate analysis of variance: a computational approach based on the Pillai-Bartlett trace. *Psychological Bulletin*, 97(1):148-152.

APPENDIX A

- Advanced Aircraft Training Climate Expert Questionnaire (AATCe-Q) -

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA



100
1908 - 2008
UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Informed consent for participation in an academic research study in the Department of Human Resource Management

THE DEVELOPMENT OF A SCALE TO MEASURE PERCEPTIONS OF THE ADVANCED AUTOMATED AIRCRAFT TRAINING CLIMATE

Research conducted by PhD student:
P. Naidoo (21346039)
Cell: +27 83 620 7299

Dear Subject Matter Expert

You are invited to participate in an academic research study because of your exceptional level of expertise and knowledge in the topic of interest, namely **advanced automated aircraft training**. This study is being conducted by Preven Naidoo (BCom, BCom Hons, MPhil, ATPL); a doctoral student in Organisational Behaviour at the University of Pretoria.

The purpose of this research survey is to establish the relevance of a number of items, questions or statements which attempt to tap the domain of a construct called **Perceptions of the Advanced Automated Aircraft Training Climate**. This questionnaire is the first phase in developing a psychometric scale to measure the training environment experiences of pilots operating advanced automated aircraft in the commercial aviation industry.

Please note:

- This study involves an anonymous survey which has been endorsed by the Airline Pilots' Association of South Africa (ALPA-SA). Your name will not appear on the questionnaire and the answers you give will be treated as strictly confidential. You cannot be identified from the answers that you give.

- By completing the questionnaire and returning it, you give your consent to participate in the study on a voluntary basis. Any data received from you will be used strictly for academic purposes and can only be accessed by the researchers.
- Your participation in this study is very important to us. Future research enhancing flight safety may depend on it. However, you may choose not to participate.
- If you do participate, please answer the questions in the attached questionnaire as completely and as honestly as possible. It should not take more than 20-30 minutes of your time to complete the questionnaire.
- This expert questionnaire consists of two parts. The first section asks for your biographical details. The second section asks for your opinion about the construct under investigation.
- The results of the study may be published in an academic journal. We will provide you with a summary of our findings on request (please supply your e-mail address on the last page for this, or send us a separate e-mail if you wish to remain anonymous).
- Please contact me, Preven (freudian@telkomsa.net) or one of my supervisors, Professor Leo Vermeulen (lvermeul@tiscali.co.za) or Professor Pieter Schaap (pieter.schaap@up.ac.za), if you have any questions or comments regarding the study. Please indicate that you have read and understand the information provided above by putting an X in this box ☐.

Expert's biographical data

Please answer the following questions to reflect the information that best represents you, by placing an X in the relevant box where applicable and answering the question or statement. This information is important in order to compile an accurate description of the panel of experts.

1. Age (years):
2. Work experience in aviation, psychology, or another relevant field (years):
3. Gender:

☐	**Female**
☐	**Male**

4. Please indicate your relevant capacity and applicable title (e.g., training captain, professor, etc.):

Capacity		**Title**
☐	Airline Pilot	
☐	Academic	
☐	Both of the above	
☐	Neither of the above	

5. Your highest academic qualification (please also specify the field of study, where applicable):

		Major field/s of specialisation
☐	Secondary School	
☐	Diploma	
☐	Bachelors	
☐	Honours	
☐	Masters	
☐	Doctorate	

6. Years of experience in <u>training</u> pilots on advanced automated aircraft, if applicable:

7. Estimated flight training experience, if applicable:

Instruction in advanced automated aircraft	**Hours**
Simulator	
Actual Aircraft	
Total instructional experience in all aircraft types	

8. Flight instructor's grade, if applicable:

9. Total flight time, if applicable (hours):

10. Please list the relevant types of aircraft you've instructed on (if applicable):

11. Please describe any pertinent information regarding your expertise which you think the researchers may find of interest (e.g., threat and error management, CRM, flight safety, flight training, accident investigation, applied psychology, etc.):

Background to the study

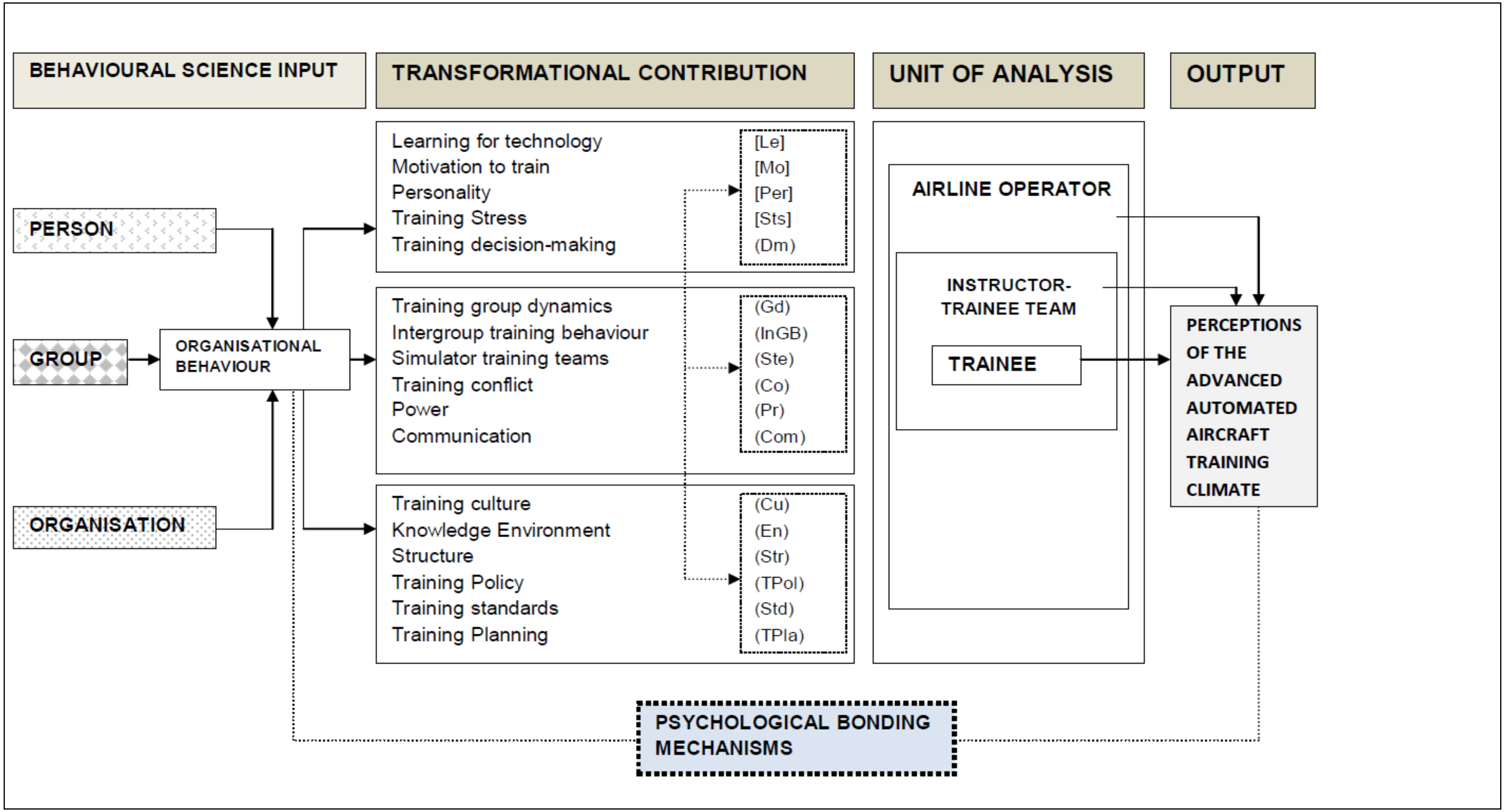
In the behavioural sciences, organisational researchers are concerned with the impact that systems have on groups and individuals. For this study, the training climate refers to 'all factors in the person, learning and organisation that influence transfer of knowledge to the job function'.

Climate must not be confused with *culture*, and the literature points out that an organisational climate refers to individuals' subjective summated sense made about understanding of policies, procedures, structure, inter- and intrapersonal constructs,. An airline pilot's perception of the training climate refers to their cognitive sense making of the psychological and organisational environment. The elements of an aviation training climate are the following:

ELEMENT LEVEL	DESCRIPTION
1. Microsphere: trainee pilot	Perceptions of learning and psychological self (academic, social, etc)
2. Mesosphere: instructional group (i.e., instructor-trainees)	Perceptions of teaching and interaction with the instructor and co-trainee
3. Macrosphere: airline operator (i.e., the organisation)	Perceptions of business atmosphere, organisational structure, policies, standards, planning, etc.

After conducting a comprehensive literature review of the relevant body of knowledge on the subject, a theoretical model of the construct (see the next figure) was developed. A list was compiled of 17 critical concepts that are important in measuring the construct. The items or statements operationalising the main construct (as used in this questionnaire) were then generated from these 17 conceptual components.

The 'super-construct' was labelled **Perceptions of the Advanced Automated Aircraft Training Climate**:


BEHAVIOURAL SCIENCE INPUT
TRANSFORMATIONAL CONTRIBUTION
UNIT OF ANALYSIS
OUTPUT
PERSON
GROUP
ORGANISATION
ORGANISATIONAL BEHAVIOUR
Learning for technology [Le]
Motivation to train [Mo]
Personality [Per]
Training Stress [Sts]
Training decision-making (Dm)
Training group dynamics (Gd)
Intergroup training behaviour (InGB)
Simulator training teams (Ste)
Training conflict (Co)
Power (Pr)
Communication (Com)
Training culture (Cu)
Knowledge Environment (En)
Structure (Str)
Training Policy (TPol)
Training standards (Std)
Training Planning (TPla)
AIRLINE OPERATOR
INSTRUCTOR-TRAINEE TEAM
TRAINEE
PERCEPTIONS OF THE ADVANCED AUTOMATED AIRCRAFT TRAINING CLIMATE
PSYCHOLOGICAL BONDING MECHANISMS

Advanced aircraft training is a combination of both simulator and route training activities. Hence, the question items or statements relate to respondents' most recent simulator and/or route training experiences where applicable, on the relevant advanced automated aircraft.

By completing this questionnaire, you will make a significant contribution to our understanding of which items in the questionnaire are essential, as opposed to ones that are useful but not essential, or not necessary at all to reflect the psychological and organisational dimensions of a modern airline pilot's training experiences. In order to ensure the integrity of the survey, it is important that you consider all the items.

There are no correct or incorrect answers.

Please consider each item individually based on your experience. Indicate your answer with an 'X' in either the 'Essential', 'Useful but not essential' or 'Not necessary' category. Also please indicate whether the item is clear or not clear to you. Please mark only **one** of the options in each case.

- If you mark a statement as **'Essential'**, this indicates that you agree that the item is strongly related to the domain and context.
- If you mark a statement as **'Useful, but not essential'**, this indicates that you consider the item to be related to the topic, but that you do not think it is important to include in the final questionnaire for scale development.
- If you mark a statement as **'Not necessary'**, this indicates that you do not think the question or statement is associated with the construct under investigation.
- If you think that an item is **not** relevant in the particular domain (person, group or organisation) where it is listed, but you feel that it is essential in one of the other domains, please mark it as "essential" and **write the letter** of the applicable domain in the last column.

Example of how to answer the questionnaire:

Number	Item statement	Consider the relevance of the item. Is it…					
		ESSENTIAL	USEFUL, BUT NOT ESSENTIAL	NOT NECESSARY			
		…in tapping the content of the construct called **Perceptions of the advanced automated aircraft training climate**?					
TRAINING CLIMATE EXPERT QUESTIONNAIRE (ORGANISATION): THE AIRLINE					**Item is clear**	**Item is not clear**	**new domain (P,G,O)**
D1	**The company's instructors are experts in the aviation industry.**	☒ essential	☐useful, but not essential	☐not necessary	☒	☐	
D2	**My company has a simulator.**	☐ essential	☒useful, but not essential	☐not necessary	☒	☐	**O**
D3	**I prefer working in a company with multi crew glass cockpit aircraft.**	☐ essential	☐useful, but not essential	☒not necessary	☒	☐	**G**
D4	**I enjoy working with computers.**	☐ essential	☐useful, but not essential	☒not necessary	☒	☐	
D5	**My company ensures that its pilots are trained in good and serviceable flight simulators.**	☒ essential	☐useful, but not essential	☐not necessary	☐	☒	

The **Advanced Aircraft Training Climate Expert Questionnaire (AATCE-Q)** starts on the next page.

Number	**Item statement**	Consider the relevance of the item. Is it…					
		ESSENTIAL	**USEFUL, BUT NOT ESSENTIAL**	**NOT NECESSARY**			
		…in tapping the content of the construct called **Perceptions of the advanced automated aircraft training climate**?					
TRAINING CLIMATE EXPERT QUESTIONNAIRE (ORGANISATION): <u>THE AIRLINE</u>					**Item clear**	**Item not clear**	**new domain (P,G,O)**
A1	**Pilot training at my airline is in line with company goals.**	☒ essential	☐useful, but not essential	☐not necessary	☒	☐	
A2	**My company's training produces world-class pilots.**	☒ essential	☐useful, but not essential	☐not necessary	☒	☐	
A3	**I have noticed a steady improvement with regard to pilot training at this company.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A4	**I know what my company's training goals are.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A5	**My company has talented people managing airline pilots' training.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A6	**Pilot training at this company is professional.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A7	**Management follows the regulator rules appropriately.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A8	**Pilot training on this aircraft is well organised at this company.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A9	**Pilots who are engaged in simulator training are professionally attired.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A10	**I understand what the company expects of me when I am in training.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A11	**It is easy to share my training experiences with colleagues at this company.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A12	**Training at my airline produces safe pilots.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A13	**There is a well-established chain of authority for pilot training on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A14	**This airline gives its pilots an appropriate amount of preparation work before training.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A15	**The paperwork involved in training for this aircraft is appropriate.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A16	**It is easy for me to appeal for assistance if I encounter a training problem at this airline.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A17	**There is sufficient training guidance from the company.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	

Number	Item statement	Consider the relevance of the item. Is it…					
		ESSENTIAL	**USEFUL, BUT NOT ESSENTIAL**	**NOT NECESSARY**			
		…in tapping the content of the construct called **Perceptions of the advanced automated aircraft training climate**?					
TRAINING CLIMATE EXPERT QUESTIONNAIRE (ORGANISATION): THE AIRLINE					**item clear**	**item not clear**	**new domain (P,G,O)**
A18	**The standard operating procedures (SOPs) for learning to fly this aircraft are adequate.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A19	**The company provided me with sufficient time to prepare for training on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A20	**The simulators my company uses to train its pilots are in good condition.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A21	**I feel motivated by my airline to train for this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A22	**The training department at my company is flexible.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A23	**The airline is very supportive of its pilots' learning requirements for this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A24	**My company's culture supports training for new technology aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A25	**There is sufficient feedback about my training on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A26	**Pilot training at my airline follows civil aviation requirements.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
A27	**My company uses only current training material.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	

Number	**Item statement**	Consider the relevance of the item. Is it…					
		ESSENTIAL	**USEFUL, BUT NOT ESSENTIAL**	**NOT NECESSARY**			
		…in tapping the content of the construct called **Perceptions of the advanced automated aircraft training climate**?					
TRAINING CLIMATE EXPERT QUESTIONNAIRE (GROUP): INSTRUCTOR-TRAINEE TEAM					**item clear**	**item not clear**	**new domain** **(P,G,O)**
B1	**I find it easy to identify with my instructor.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B2	**I can easily identify with my simulator partner.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B3	**I work well with others during simulator training exercises.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B4	**Instructors communicate their expectations effectively.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B5	**I learn better when I work as a member of the crew.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B6	**I am always at ease when interacting with my flight instructor.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B7	**I always find my simulator partner prepared for training.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B8	**I trust my simulator partner.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B9	**I am confident that my instructor will be fair.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B10	**I operate well as a crew member in the simulator.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B11	**My instructor is willing to listen.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B12	**I communicate well with my simulator partner.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B13	**I feel secure in the decisions made by my simulator partner.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B14	**I make good decisions with my partner in the simulator.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B15	**I find that decision-making with my simulator partner is equitable.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B16	**I am motivated by my instructor.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B17	**When training for this aircraft, I feel that I am part of a team.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	

Number	**Item statement**	Consider the relevance of the item. Is it…					
		ESSENTIAL	**USEFUL, BUT NOT ESSENTIAL**	**NOT NECESSARY**			
		…in tapping the content of the construct called **Perceptions of the advanced automated aircraft training climate**?					
TRAINING CLIMATE EXPERT QUESTIONNAIRE (GROUP): INSTRUCTOR-TRAINEE TEAM					**item clear**	**item not clear**	**new domain (P,G,O)**
B18	**The instructors on this aircraft are committed.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B19	**Instructors are similar in how they teach pilots to fly this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B20	**I am always paired with someone who is committed to performing well.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B21	**I enjoy being evaluated as a member of a crew.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B22	**Instructors on this fleet follow company policy.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B23	**The instructors on this aircraft avoid overloading pilots with unnecessary information.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B24	**I always bond well with my simulator partner.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B25	**Decisions made in flight simulator training exercises are team-based.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B26	**The instructors on this aircraft are friendly.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
B27	**I get sufficient feedback on my flight training performance.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	

Number	**Item statement**	Consider the relevance of the item. Is it…					
		ESSENTIAL	**USEFUL, BUT NOT ESSENTIAL**	**NOT NECESSARY**			
		…in tapping the content of the construct called **Perceptions of the advanced automated aircraft training climate**?					
TRAINING CLIMATE EXPERT QUESTIONNAIRE (INDIVIDUAL): TRAINEE					**item clear**	**item not clear**	**new domain** **(P,G,O)**
C1	**Pilots are in direct control of the training outcome.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C2	**A good training session on this aircraft is a result of the trainee's actions.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C3	**Evaluation of my flight training is objective.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C4	**Adequate preparation improves flight training performance.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C5	**I am always on time for a flight training session.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C6	**I co-operate well when training in a simulator.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C7	**I never feel rushed in the flight simulator.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C8	**I easily express my opinion during flight training.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C9	**I prepare sufficiently for training on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C10	**After flight training, I feel a sense of mastery.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C11	**I enjoy learning about this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C12	**Simulator training affects behaviour on the actual aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C13	**I get along well with my flight simulator partners.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C14	**I found my transition to advanced automated aircraft easy.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C15	**I believe that if pilots do well in training, overall flight safety improves.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C16	**I am happy with simulator training on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C17	**I aim to do better at my next flight simulator training session by learning from my mistakes.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	

Number	**Item statement**	Consider the relevance of the item. Is it…					
		ESSENTIAL	**USEFUL, BUT NOT ESSENTIAL**	**NOT NECESSARY**			
		…in tapping the content of the construct called **Perceptions of the advanced automated aircraft training climate**?					
TRAINING CLIMATE EXPERT QUESTIONNAIRE (INDIVIDUAL): TRAINEE					**item clear**	**item not clear**	**new domain (P,G,O)**
C18	**I have a positive relationship with my colleagues.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C19	**The workload between trainees is balanced during a flight simulator training session.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C20	**Pilots are judged as members of a team when they train in the flight simulator.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C21	**I feel rewarded for the amount of work I put into flight training.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C22	**The more work I put into my preparation for training on this aircraft, the better I will perform.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C23	**Pilots who are prepared have no problems training for this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C24	**It is essential that pilots prepare adequately to pass a rating on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C25	**I am in control of the outcome of my flight training on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C26	**I enjoy studying the technical aspects of the aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C27	**I always learn something new after undergoing training on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C28	**I focus on the pertinent and relevant topics when learning about this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C29	**I reflect on my learning after a flight training experience.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C30	**I look for additional information so as to gain a deeper understanding of this aircraft's systems.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C31	**I know where to find specific information for this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C32	**It is important to know more than just what is required to pass.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	

Number	**Item statement**	Consider the relevance of the item. Is it…					
		ESSENTIAL	**USEFUL, BUT NOT ESSENTIAL**	**NOT NECESSARY**			
		…in tapping the content of the construct called **Perceptions of the advanced automated aircraft training climate**?					
TRAINING CLIMATE EXPERT QUESTIONNAIRE (INDIVIDUAL): TRAINEE					**item clear**	**item not clear**	**new domain (P,G,O)**
C33	**I find the training on this aircraft easy.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C34	**I am relaxed in the flight simulator.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C35	**I find the training on this aircraft easy.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C36	**I do well in training for this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C37	**I look forward to my next flight training session.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C38	**I sleep well the night before training on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C39	**An appropriate level of stress helps me perform well in flight training for this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C40	**I'm comfortable undergoing training for this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C41	**I can control my anxiety so as to perform well in training.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C42	**I enjoy spending extra time flight training.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C43	**I am motivated to learn more about this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C44	**I am happy to be subjected to regular flight checks.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C45	**I enjoy route training on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C46	**I enjoy simulator training for this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C47	**If my simulator partner is having a bad day, I am not affected.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C48	**I create a relaxed atmosphere in the flight simulator.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	

C49	**The length of time spent simulator training is appropriate for this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C50	**I enjoy the free play flight simulator time on this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C51	**I aim to gain a deeper understanding of this aircraft.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	
C52	**I learn more than the company requires me to.**	☐ essential	☐useful, but not essential	☐not necessary	☐	☐	

1. Clarity

Aviation psychometric measurement items should be well written, distinct, and at an appropriate reading level for professional pilots employed on various types of advanced automated aircraft from airline organisations (in the private and the public sector), from a diverse population.

Please list any items that, in your opinion, are not clear. Also provide suggestions on how to clarify these items.

2. Comprehensiveness

Do you think that the dimensions (statements/questions) of the desired content behavioural domains are adequate in tapping the construct?

Please suggest which items should be deleted; or provide additional/alternative items you think may be relevant.

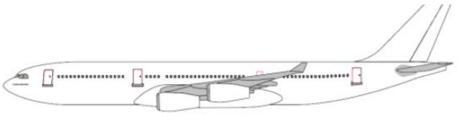
3. Please provide us with your email address if you would like to receive feedback:

Thank you for your time and participation. <u>Please save this document then e-mail the completed questionnaire to:</u> **freudian@telkomsa.net**

APPENDIX B

- Survey Invitation Letter-

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA



100
1908 - 2008
UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Participation in an academic research study

Advanced Aircraft Training Climate Questionnaire (AATC-Q)

UNIVERSITY OF PRETORIA
Faculty of Economic and Management Sciences (EMS)
Department of Human Resource Management
Telephone: 012 420 3074

Dear Colleague

You are invited to participate in an academic research study because of your experience and knowledge in the research area, namely **advanced automated aircraft training**. This study is being conducted by the unit of Organisational Behaviour at the University of Pretoria.

The purpose of this study is to explore your perceptions and experiences regarding training to operate advanced automated aircraft. The information obtained from this project will add to the current body of knowledge on human factors, automation, pilot training and flight safety. Your co-operation in completing the questionnaire will be a valuable input to the overall success of the study.

Please note:

> This study involves an anonymous survey which has been endorsed by the Airline Pilots' Association of South Africa (ALPA-SA). Your name will not appear on the questionnaire and the answers you give will be treated as strictly confidential. Furthermore, you cannot be identified from the answers that you give.
>
> By completing the questionnaire and returning it, you give your consent to participate in the study on a voluntary basis. Any data received from you will be used strictly for academic purposes and can only be accessed by the researchers.

Your participation in this study is very important to us. Future research enhancing flight safety may depend on it. However, you may choose not to participate.

If you do participate, please answer the questions in the attached questionnaire as completely and as honestly as possible. It should not take more than 20-30 minutes of your time to complete the questionnaire.

This questionnaire consists of three parts. The first section asks for your demographical details. The second section asks for your opinion on a statement. The third part is reserved for your feedback.

The results of the study may be published in an academic journal. We will provide you with a summary of our findings on request (please supply your e-mail address on the last page for this, or send us a separate e-mail if you wish to remain anonymous).

Please contact one of the researchers directly if you have any questions or comments regarding the study:

- Professor Leo Vermeulen (lvermeul@tiscali.co.za). (Supervisor).
- Professor Pieter Schaap (pieter.schaap@up.ac.za). (Co-supervisor).
- Preven Naidoo (freudian@telkomsa.net), 083 620 7299. (Research student).

Please indicate that you have read the information provided above by putting an X in this box ☐.

Thank you for your time and participation

Yours sincerely

Professor Leo Vermeulen

Unit of Aviation Management

UNIVERSITY OF PRETORIA

APPENDIX C

- Three Scale Items -

SCALE 1: ORGANISATIONAL PROFESSIONALISM

Training on this aircraft is well organised.
Training on this aircraft is professional.
My company’s training produces world class pilots.
Training at my airline is in line with company goals.
The airline is very supportive of its pilots’ learning requirements for this aircraft.
There is sufficient training guidance from the company.
Management follows the rules and regulations appropriately.
My company’s culture supports training for new technology aircraft.
I understand what the company expects of me when training.
My company has talented people in training.
If I had to experience a problem in training, it’s easy for me to appeal.
I know what my company’s training goals are.
Training at my airline produces safe pilots.
There is sufficient feedback about my training on this aircraft.
My company uses only current training material.
Training is in line with civil aviation regulations.
The airline gives its pilots an appropriate amount of preparation work for training.
My instructor is willing to listen.
Pilots are in direct control of the training outcome.
I’m given sufficient time to prepare for training on this aircraft.

SCALE 2: INTRINSIC MOTIVATION

It’s a good idea to know more than what is required.
I try never to be late for a training session.
I co-operate when training in a simulator.
I aim to gain a deeper understanding of this aircraft.
Preparation improves performance.
I read to understand so as to gain a deeper understanding of this aircraft’s systems.
I have a positive relationship with my colleagues.
I operate well as a crew member in the simulator.
I enjoy studying the technical aspects of the aircraft.

SCALE 3: INDIVIDUAL CONTROL OF TRAINING OUTCOMES

I’m comfortable undergoing training for this aircraft.
I’m in control of the outcome of a training session.
I can control my anxiety so as to perform well in training.
The instructors on this aircraft don’t overload us with information.

APPENDIX D

- Informed consent form -

Faculty of Economic and
Management Sciences

Informed consent for participation in an academic research study

Dept. of Human Resource Management

THE DEVELOPMENT OF A SCALE TO MEASURE PERCEPTIONS OF THE ADVANCED AUTOMATED AIRCRAFT TRAINING CLIMATE

Research conducted by:
P. Naidoo (21346039)
Cell: 083 620 7299

Dear Participant

You are invited to participate in an academic research study being conducted by Preven Naidoo (BCom AVM, BCom Hons BM, MPhil HRM, ATPL), a Doctoral student in Organisational Behaviour at the University of Pretoria's unit for Aviation Management. The purpose of the study is to develop a psychological measurement of airline pilots' perceptions of their training environment, specifically associated with advanced automated aircraft and its related systems.

Please note the following:

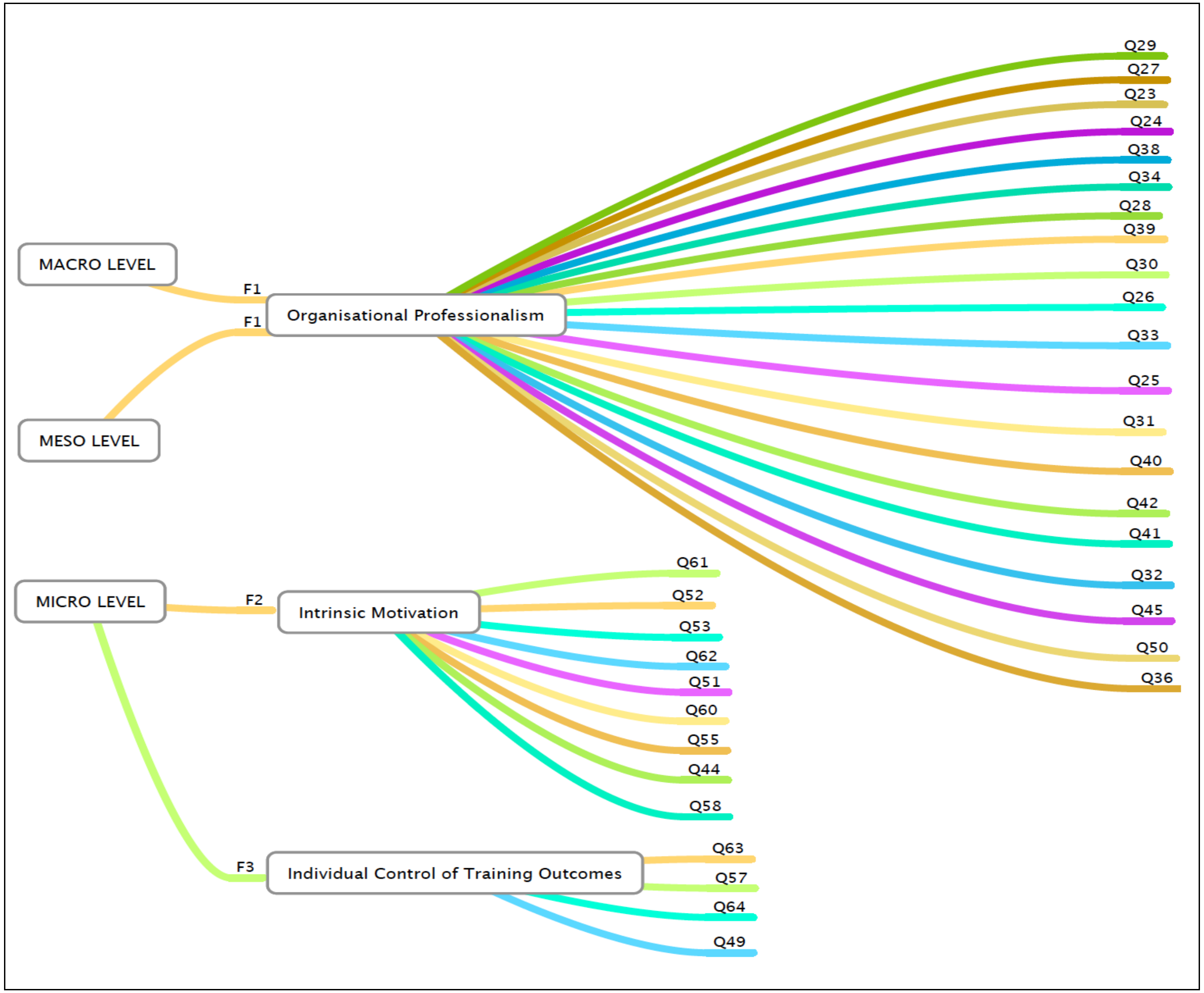
- This study involves an anonymous survey. Your name will not appear on the questionnaire and the answers you give will be treated as strictly confidential. You cannot be identified in person based on the answers you give. Note also that the study has been scrutinised and passed by the University's ethics committee.
- By completing the questionnaire and returning it, you are giving your consent to participate in the study on a voluntary basis. Furthermore, all data received by you will be used for academic purposes only and can only be accessed by the researchers.
- Your participation in this study is very important to us and future research for enhancing flight safety. You may however, choose not to participate.
- Please answer the questions found in the attached questionnaire as completely and honestly as possible. This should not take more than 20 minutes of your time.
- The results of the study will be used for academic purposes only and may be published in a scientific journal. We will provide you with a summary of our findings on request.
- Please contact me (freudian@telkomsa.net) or my supervisor, Professor Leo Vermeulen (lvermeul@tiscali.co.za) if you have any questions, comments or additional information regarding the study.
- Please indicate that you have read and understand the information provided above by ticking this box☐.
-

Kindest regards,

Professor Leo Vermeulen (Unit for Aviation Research, University of Pretoria).

APPENDIX E

- Illustrated structure of the measurement construct -


MACRO LEVEL
F1
MESO LEVEL
F1
Organisational Professionalism
Q29
Q27
Q23
Q24
Q38
Q34
Q28
Q39
Q30
Q26
Q33
Q25
Q31
Q40
Q42
Q41
Q32
Q45
Q50
Q36
MICRO LEVEL
F2
Intrinsic Motivation
Q61
Q52
Q53
Q62
Q51
Q60
Q55
Q44
Q58
F3
Individual Control of Training Outcomes
Q63
Q57
Q64
Q49

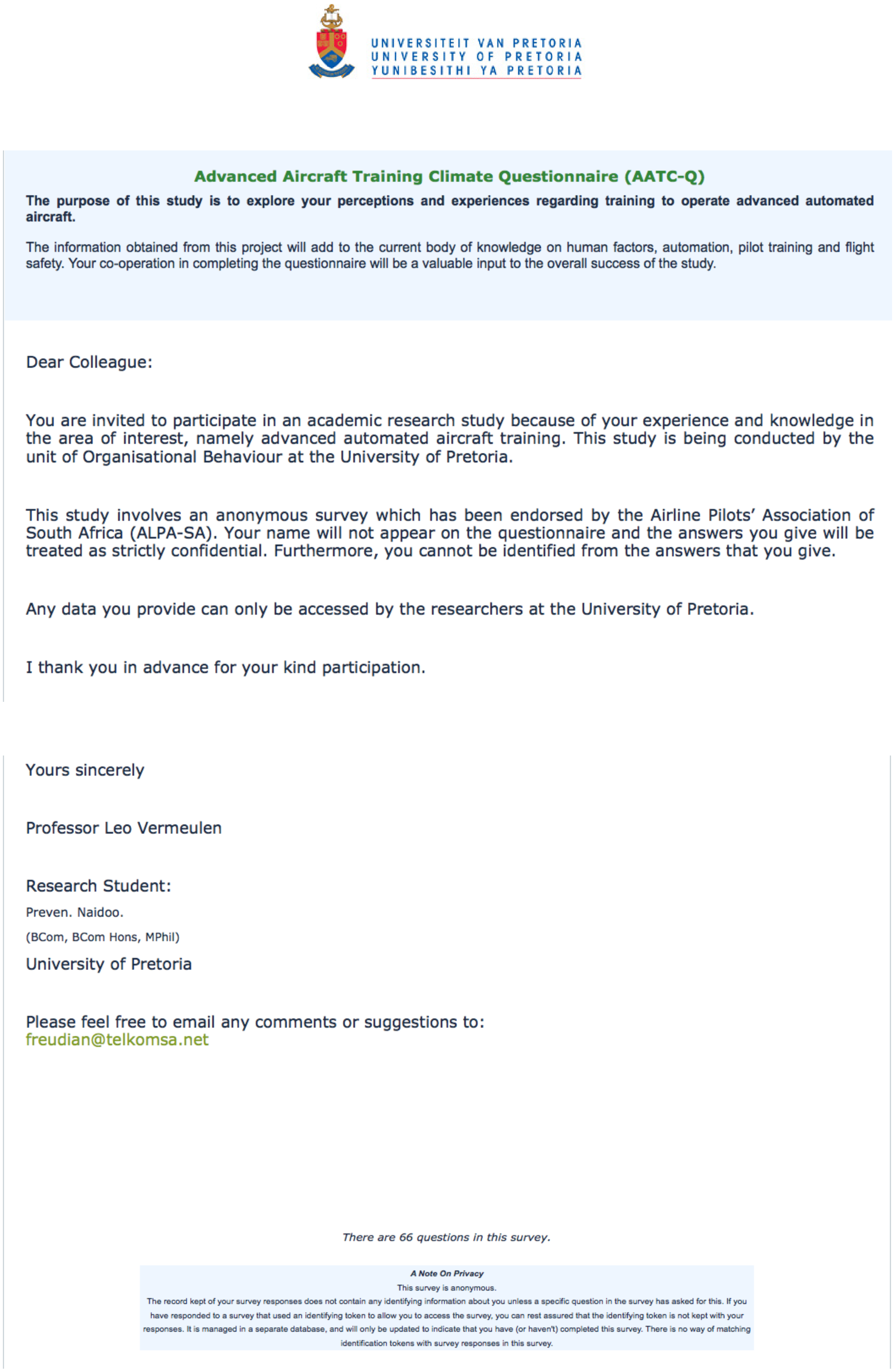
APPENDIX F

- Web based survey -

Note about the web-based version of the large-scale survey:

Because the web-based survey can accommodate items in specific format only, the survey items were adapted to fit such categories. Nonetheless, the essence and content of an item was not changed. Furthermore the survey software (Lime Survey) requires that, all surveys must have at least one group. Groups are used to 'group' questions/items together into logical categories. A group has a name and a description. Each item of the web survey was mandatory, thus an asterisk marking the item/question.

The following screen shots provide an illustration of the appearance of the actual web-based survey. The survey was accessed via an internet hyperlink.



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Advanced Aircraft Training Climate Questionnaire (AATC-Q)

The purpose of this study is to explore your perceptions and experiences regarding training to operate advanced automated aircraft.

The information obtained from this project will add to the current body of knowledge on human factors, automation, pilot training and flight safety. Your co-operation in completing the questionnaire will be a valuable input to the overall success of the study.

Dear Colleague:

You are invited to participate in an academic research study because of your experience and knowledge in the area of interest, namely advanced automated aircraft training. This study is being conducted by the unit of Organisational Behaviour at the University of Pretoria.

This study involves an anonymous survey which has been endorsed by the Airline Pilots' Association of South Africa (ALPA-SA). Your name will not appear on the questionnaire and the answers you give will be treated as strictly confidential. Furthermore, you cannot be identified from the answers that you give.

Any data you provide can only be accessed by the researchers at the University of Pretoria.

I thank you in advance for your kind participation.

Yours sincerely

Professor Leo Vermeulen

Research Student:

Preven. Naidoo.

(BCom, BCom Hons, MPhil)

University of Pretoria

Please feel free to email any comments or suggestions to:
freudian@telkomsa.net

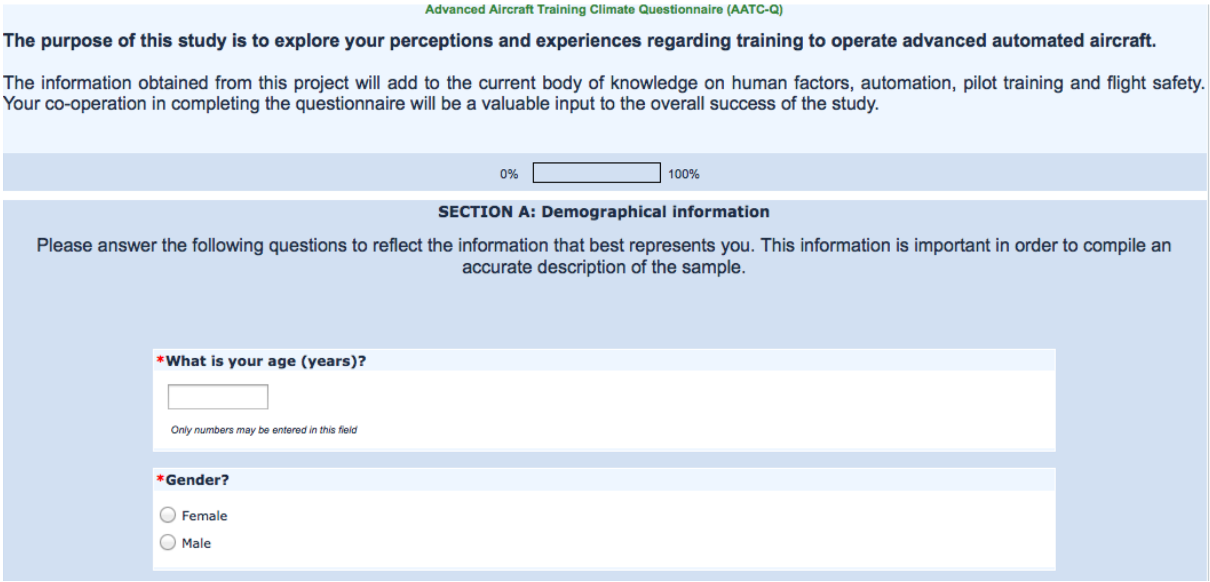
There are 66 questions in this survey.

A Note On Privacy

This survey is anonymous.

The record kept of your survey responses does not contain any identifying information about you unless a specific question in the survey has asked for this. If you have responded to a survey that used an identifying token to allow you to access the survey, you can rest assured that the identifying token is not kept with your responses. It is managed in a separate database, and will only be updated to indicate that you have (or haven't) completed this survey. There is no way of matching identification tokens with survey responses in this survey.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA



Advanced Aircraft Training Climate Questionnaire (AATC-Q)

The purpose of this study is to explore your perceptions and experiences regarding training to operate advanced automated aircraft.

The information obtained from this project will add to the current body of knowledge on human factors, automation, pilot training and flight safety. Your co-operation in completing the questionnaire will be a valuable input to the overall success of the study.

0% 100%

SECTION A: Demographical information

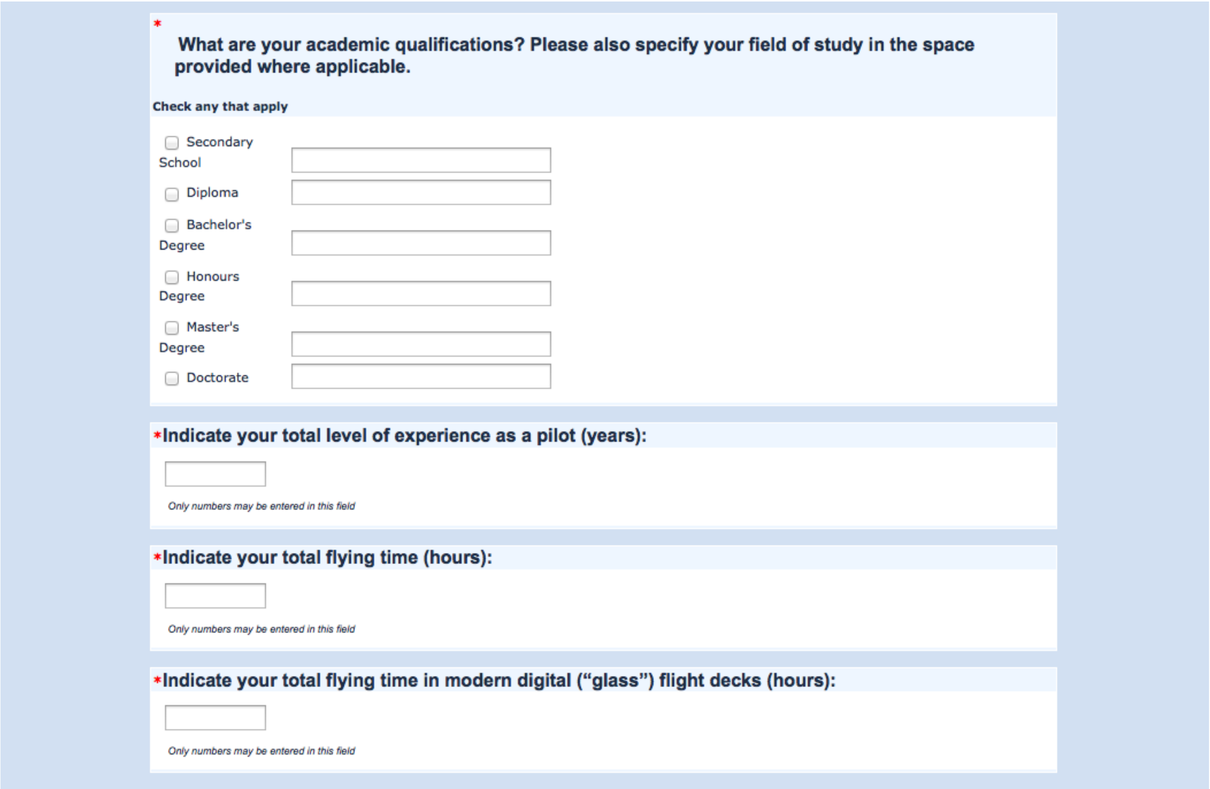
Please answer the following questions to reflect the information that best represents you. This information is important in order to compile an accurate description of the sample.

***What is your age (years)?**

Only numbers may be entered in this field

***Gender?**

○ Female
○ Male



*** What are your academic qualifications? Please also specify your field of study in the space provided where applicable.**

Check any that apply

☐ Secondary School
☐ Diploma
☐ Bachelor's Degree
☐ Honours Degree
☐ Master's Degree
☐ Doctorate

***Indicate your total level of experience as a pilot (years):**

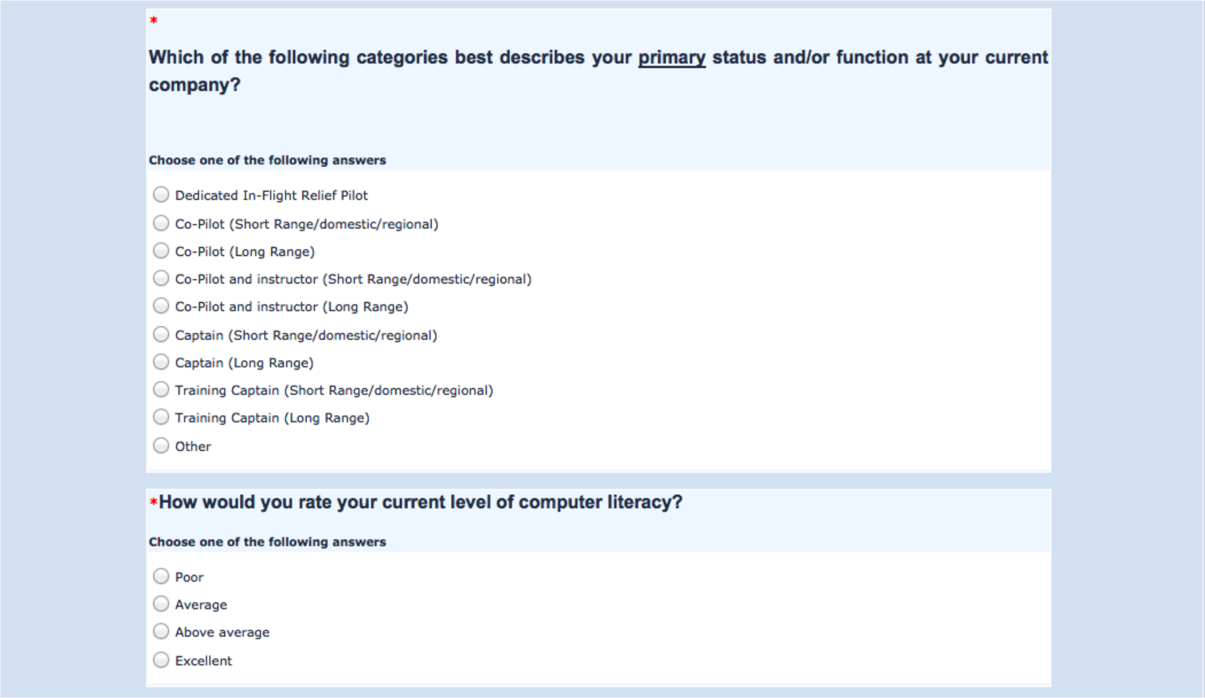
Only numbers may be entered in this field

***Indicate your total flying time (hours):**

Only numbers may be entered in this field

***Indicate your total flying time in modern digital (“glass”) flight decks (hours):**

Only numbers may be entered in this field



*

Which of the following categories best describes your primary status and/or function at your current company?

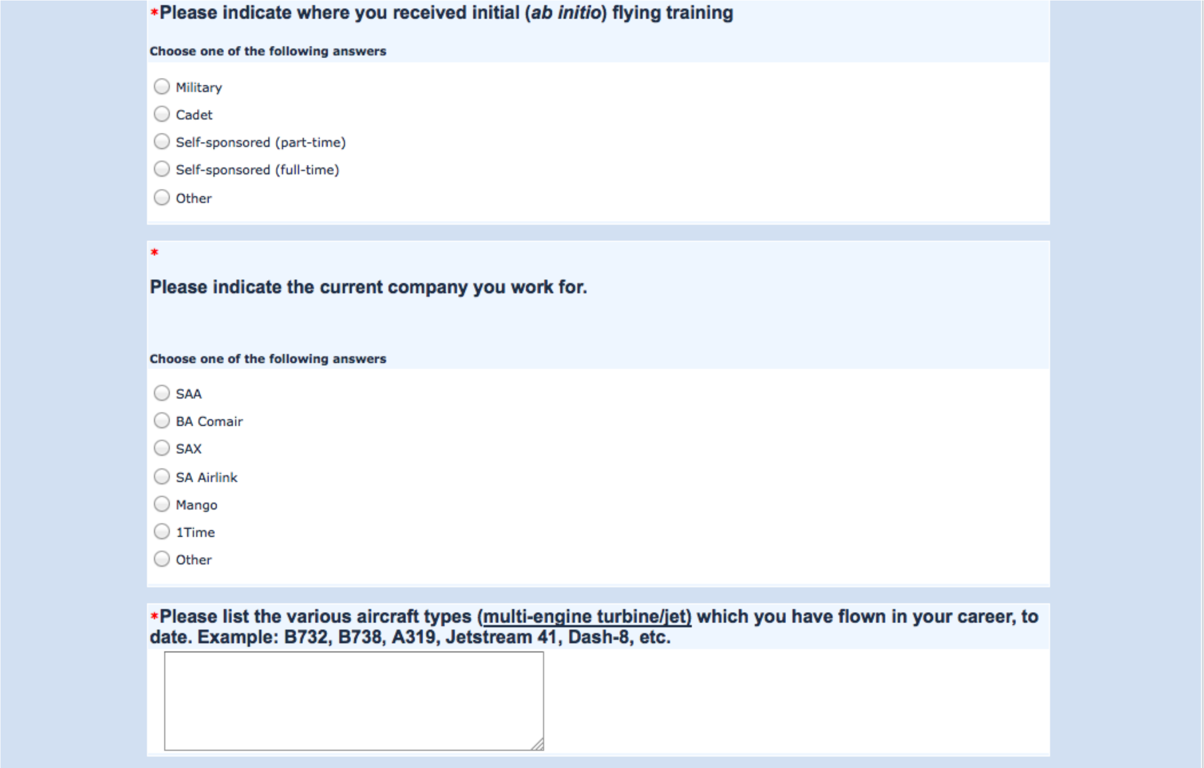
Choose one of the following answers

○ Dedicated In-Flight Relief Pilot
○ Co-Pilot (Short Range/domestic/regional)
○ Co-Pilot (Long Range)
○ Co-Pilot and instructor (Short Range/domestic/regional)
○ Co-Pilot and instructor (Long Range)
○ Captain (Short Range/domestic/regional)
○ Captain (Long Range)
○ Training Captain (Short Range/domestic/regional)
○ Training Captain (Long Range)
○ Other

***How would you rate your current level of computer literacy?**

Choose one of the following answers

○ Poor
○ Average
○ Above average
○ Excellent



***Please indicate where you received initial (*ab initio*) flying training**

Choose one of the following answers

○ Military
○ Cadet
○ Self-sponsored (part-time)
○ Self-sponsored (full-time)
○ Other

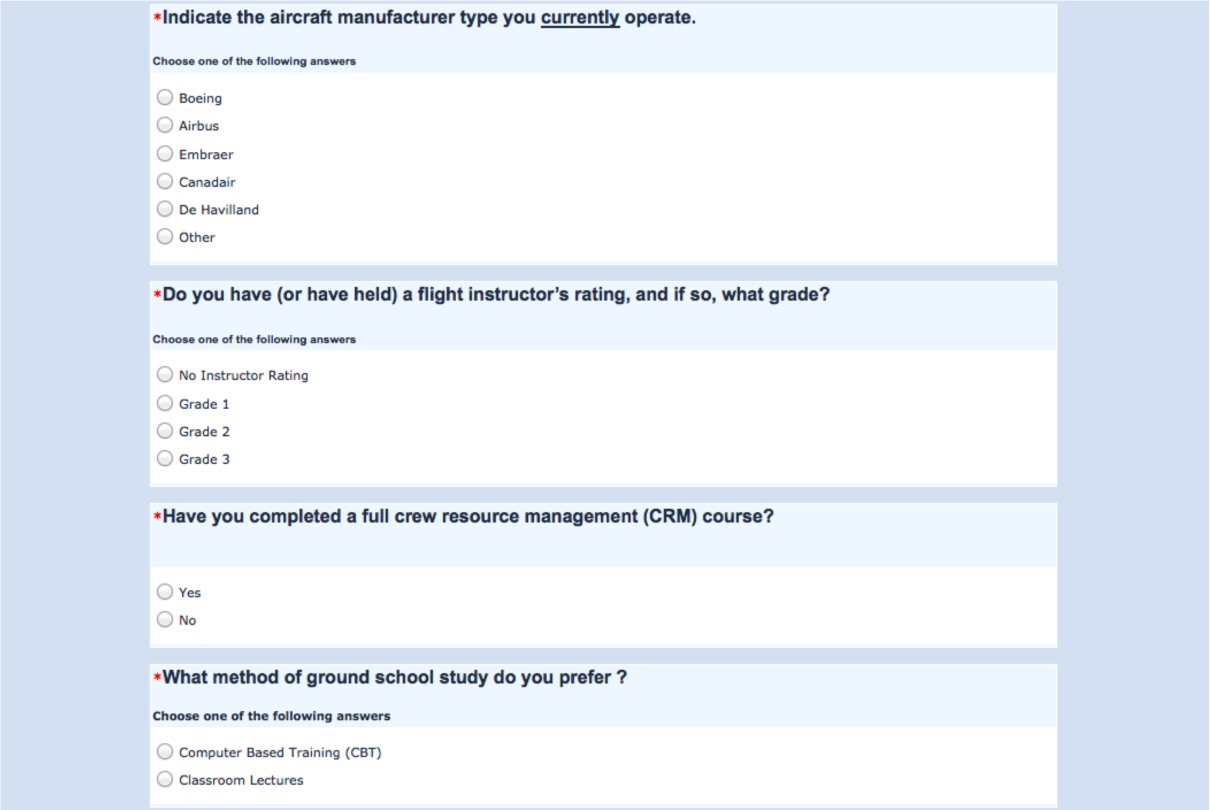
*

Please indicate the current company you work for.

Choose one of the following answers

○ SAA
○ BA Comair
○ SAX
○ SA Airlink
○ Mango
○ 1Time
○ Other

***Please list the various aircraft types (multi-engine turbine/jet) which you have flown in your career, to date. Example: B732, B738, A319, Jetstream 41, Dash-8, etc.**



***Indicate the aircraft manufacturer type you currently operate.**

Choose one of the following answers

- ○ Boeing
- ○ Airbus
- ○ Embraer
- ○ Canadair
- ○ De Havilland
- ○ Other

***Do you have (or have held) a flight instructor's rating, and if so, what grade?**

Choose one of the following answers

- ○ No Instructor Rating
- ○ Grade 1
- ○ Grade 2
- ○ Grade 3

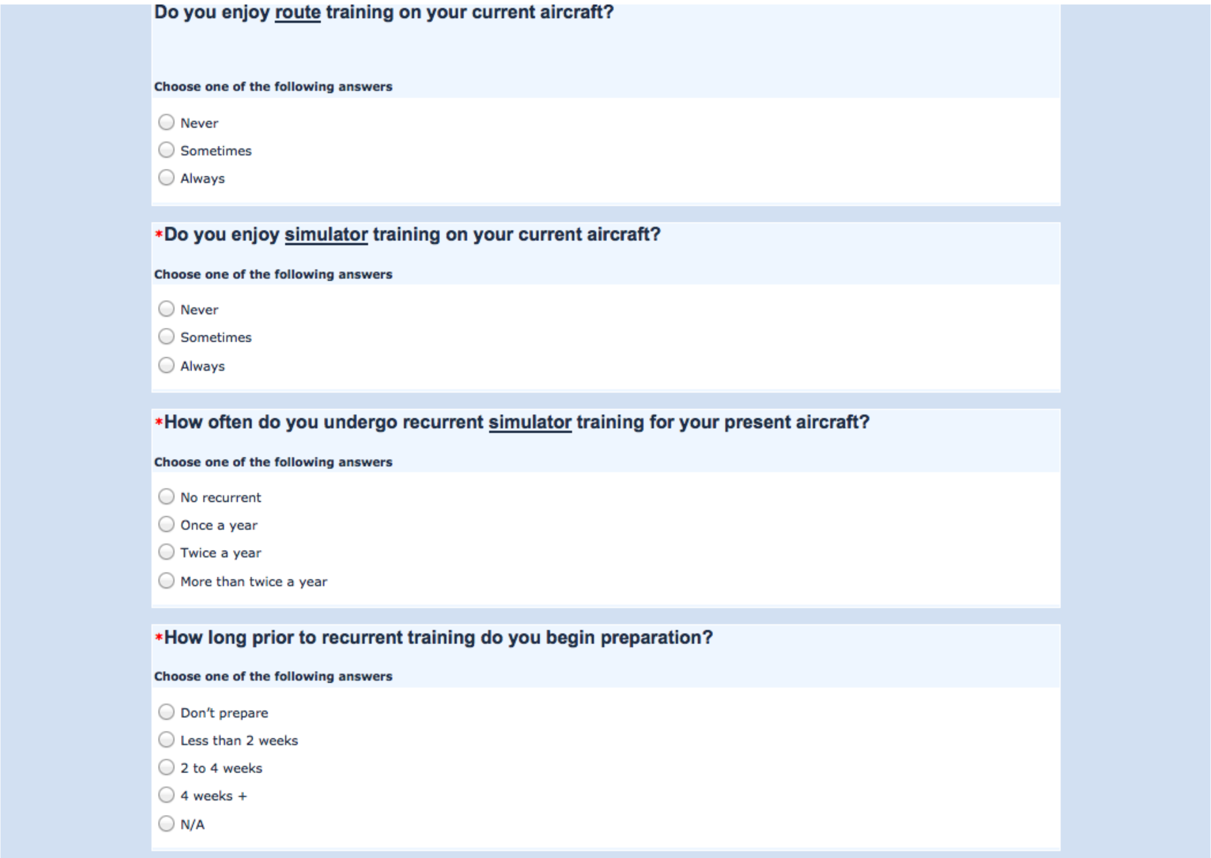
***Have you completed a full crew resource management (CRM) course?**

- ○ Yes
- ○ No

***What method of ground school study do you prefer ?**

Choose one of the following answers

- ○ Computer Based Training (CBT)
- ○ Classroom Lectures



Do you enjoy route training on your current aircraft?

Choose one of the following answers

- ○ Never
- ○ Sometimes
- ○ Always

***Do you enjoy simulator training on your current aircraft?**

Choose one of the following answers

- ○ Never
- ○ Sometimes
- ○ Always

***How often do you undergo recurrent simulator training for your present aircraft?**

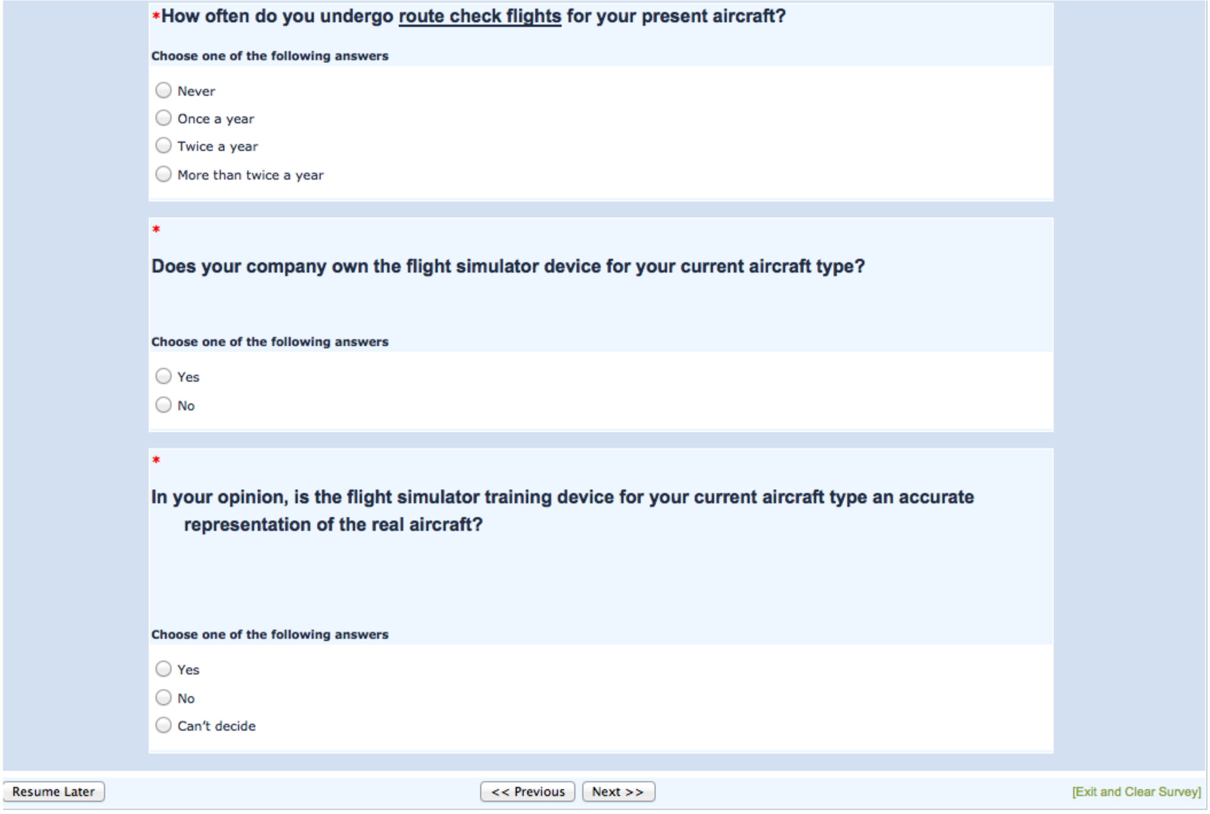
Choose one of the following answers

- ○ No recurrent
- ○ Once a year
- ○ Twice a year
- ○ More than twice a year

***How long prior to recurrent training do you begin preparation?**

Choose one of the following answers

- ○ Don't prepare
- ○ Less than 2 weeks
- ○ 2 to 4 weeks
- ○ 4 weeks +
- ○ N/A



***How often do you undergo route check flights for your present aircraft?**

Choose one of the following answers

○ Never
○ Once a year
○ Twice a year
○ More than twice a year

*

Does your company own the flight simulator device for your current aircraft type?

Choose one of the following answers

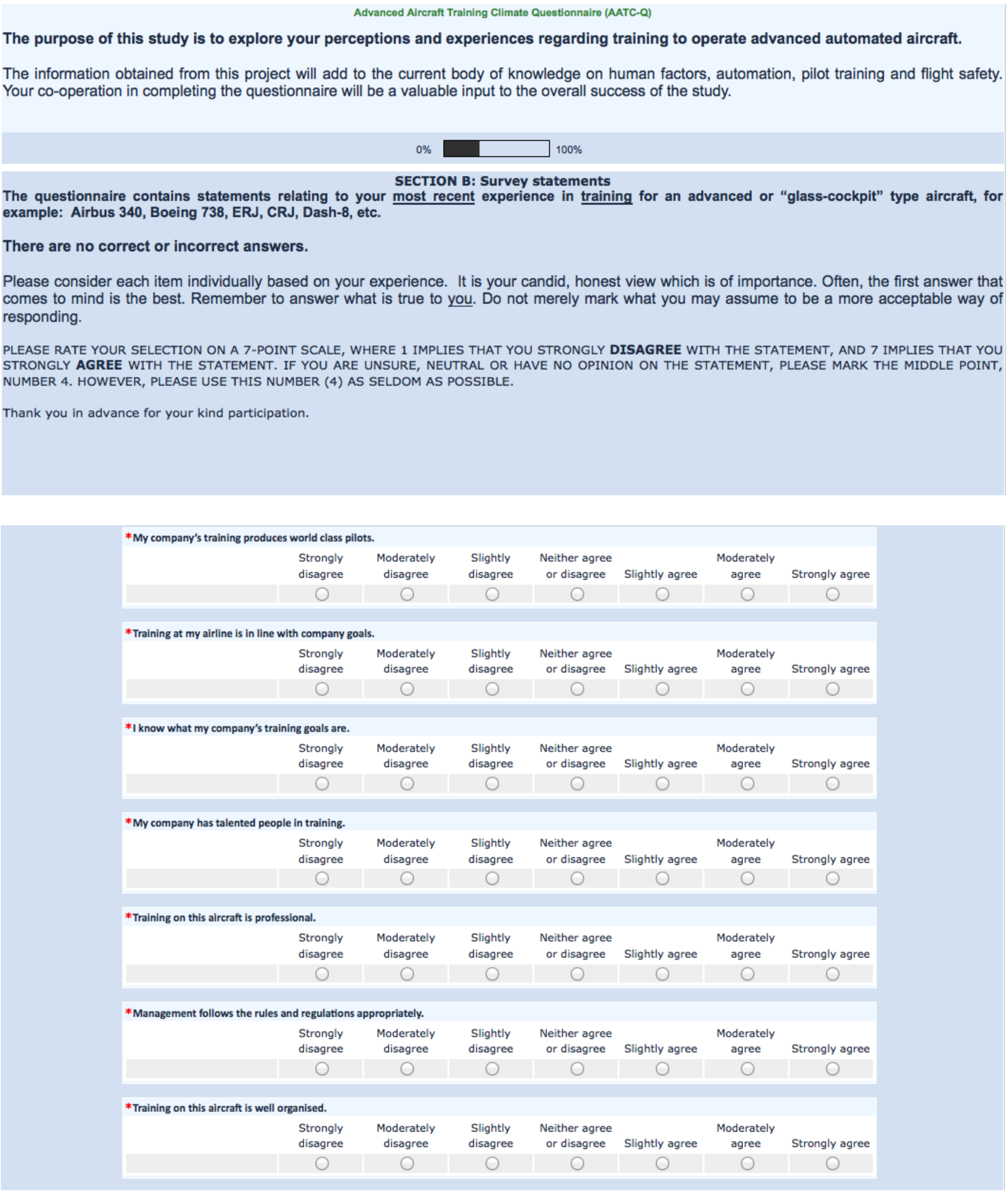
○ Yes
○ No

*

In your opinion, is the flight simulator training device for your current aircraft type an accurate representation of the real aircraft?

Choose one of the following answers

○ Yes
○ No
○ Can't decide



Advanced Aircraft Training Climate Questionnaire (AATC-Q)

The purpose of this study is to explore your perceptions and experiences regarding training to operate advanced automated aircraft.

The information obtained from this project will add to the current body of knowledge on human factors, automation, pilot training and flight safety. Your co-operation in completing the questionnaire will be a valuable input to the overall success of the study.

0% ▬▭▭ 100%

SECTION B: Survey statements

The questionnaire contains statements relating to your most recent experience in training for an advanced or "glass-cockpit" type aircraft, for example: Airbus 340, Boeing 738, ERJ, CRJ, Dash-8, etc.

There are no correct or incorrect answers.

Please consider each item individually based on your experience. It is your candid, honest view which is of importance. Often, the first answer that comes to mind is the best. Remember to answer what is true to you. Do not merely mark what you may assume to be a more acceptable way of responding.

PLEASE RATE YOUR SELECTION ON A 7-POINT SCALE, WHERE 1 IMPLIES THAT YOU STRONGLY **DISAGREE** WITH THE STATEMENT, AND 7 IMPLIES THAT YOU STRONGLY **AGREE** WITH THE STATEMENT. IF YOU ARE UNSURE, NEUTRAL OR HAVE NO OPINION ON THE STATEMENT, PLEASE MARK THE MIDDLE POINT, NUMBER 4. HOWEVER, PLEASE USE THIS NUMBER (4) AS SELDOM AS POSSIBLE.

Thank you in advance for your kind participation.

***My company's training produces world class pilots.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***Training at my airline is in line with company goals.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***I know what my company's training goals are.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***My company has talented people in training.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***Training on this aircraft is professional.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***Management follows the rules and regulations appropriately.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***Training on this aircraft is well organised.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*I understand what the company expects of me when training.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*Training at my airline produces safe pilots.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*The airline gives its pilots an appropriate amount of preparation work for training.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*If I had to experience a problem in training, it's easy for me to appeal.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*There is sufficient training guidance from the company.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*The standard operating procedures (SOPs) for learning to fly this aircraft is adequate.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*I'm given sufficient time to prepare for training on this aircraft.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*The simulators my company trains its pilots in are in good condition.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*The airline is very supportive of its pilots' learning requirements for this aircraft.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*My company's culture supports training for new technology aircraft.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*There is sufficient feedback about my training on this aircraft.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*Training is in line with civil aviation regulations.

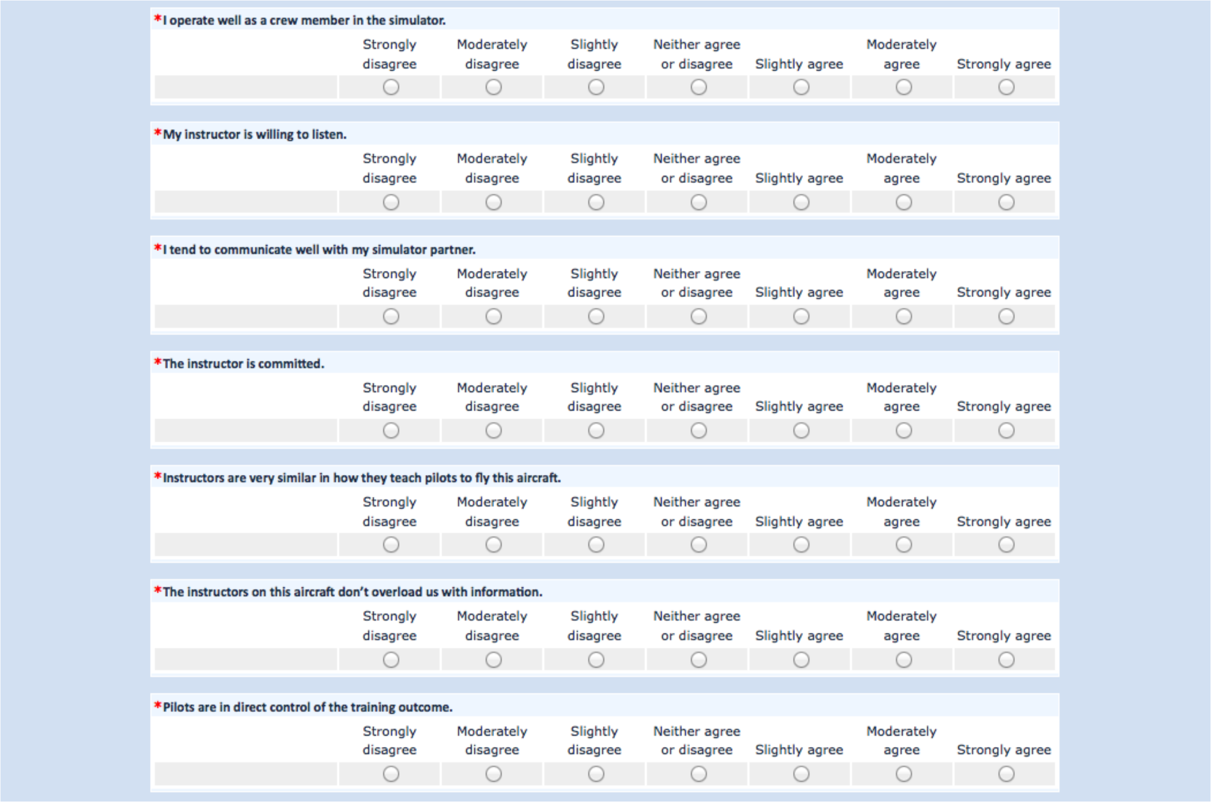
	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*My company uses only current training material.

	Strongly Disagree	Slightly Disagree	Disagree	Neither Disagree, nor Agree	Agree	Slightly Agree	Strongly Agree
	○	○	○	○	○	○	○

*I learn better when I work as a member of the crew.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○



***I operate well as a crew member in the simulator.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***My instructor is willing to listen.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***I tend to communicate well with my simulator partner.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***The instructor is committed.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***Instructors are very similar in how they teach pilots to fly this aircraft.**

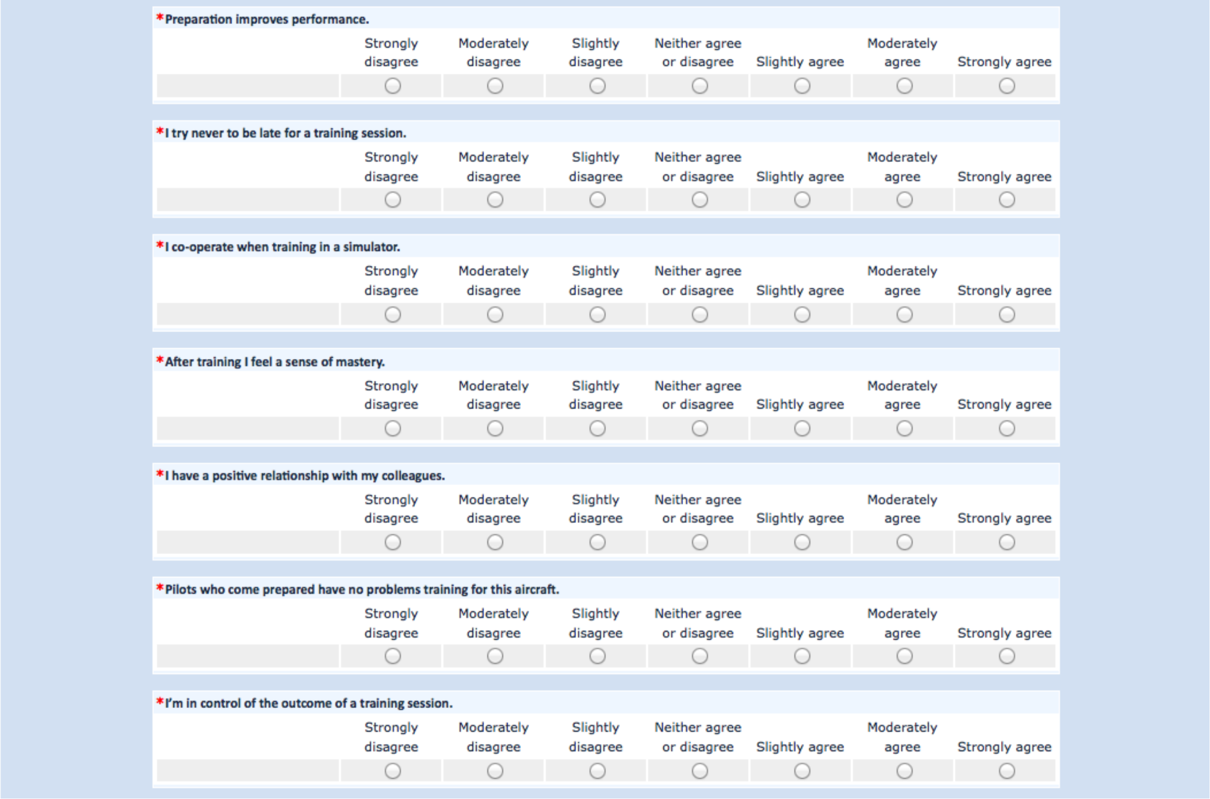
	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***The instructors on this aircraft don't overload us with information.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***Pilots are in direct control of the training outcome.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○



***Preparation improves performance.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***I try never to be late for a training session.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***I co-operate when training in a simulator.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***After training I feel a sense of mastery.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***I have a positive relationship with my colleagues.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***Pilots who come prepared have no problems training for this aircraft.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

***I'm in control of the outcome of a training session.**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
	○	○	○	○	○	○	○

*I enjoy studying the technical aspects of the aircraft.

Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

*I reflect on my learning experience after a simulator session.

Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

*I read to understand so as to gain a deeper understanding of this aircraft's systems.

Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

*It's a good idea to know more than what is required.

Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

*I aim to gain a deeper understanding of this aircraft.

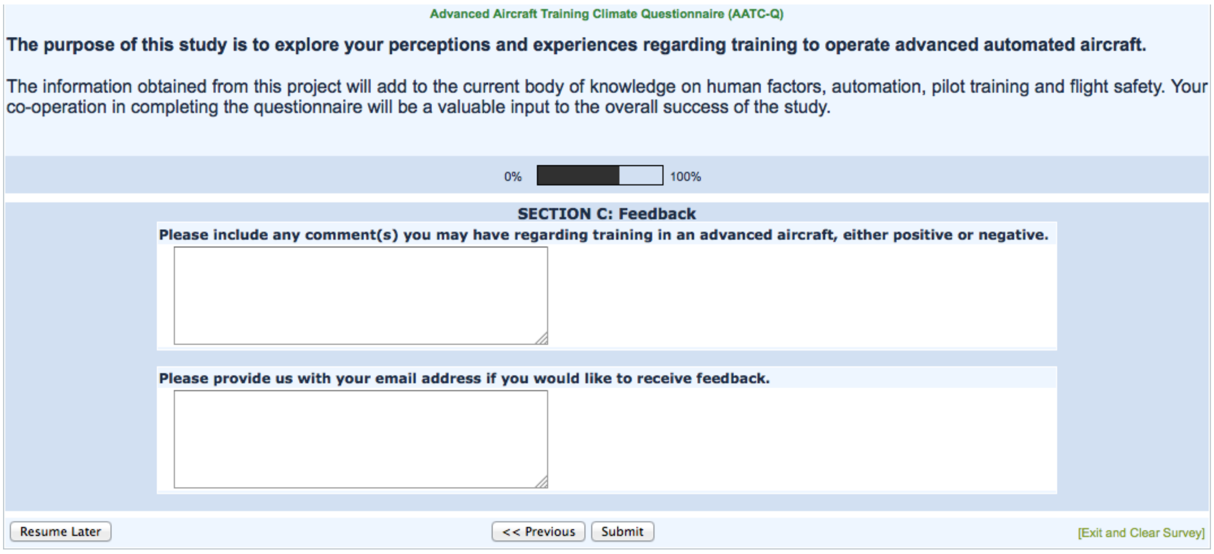
Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

*I'm comfortable undergoing training for this aircraft.

Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

*I can control my anxiety so as to perform well in training.

Strongly disagree	Moderately disagree	Slightly disagree	Neither agree or disagree	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○



Advanced Aircraft Training Climate Questionnaire (AATC-Q)

The purpose of this study is to explore your perceptions and experiences regarding training to operate advanced automated aircraft.

The information obtained from this project will add to the current body of knowledge on human factors, automation, pilot training and flight safety. Your co-operation in completing the questionnaire will be a valuable input to the overall success of the study.

SECTION C: Feedback

Please include any comment(s) you may have regarding training in an advanced aircraft, either positive or negative.

Please provide us with your email address if you would like to receive feedback.